Quantitative Analysis of the Proteomic Response of Human Tumor Cell Lines to Sirtinol

Thesis

Quantitative Analysis of the Proteomic Response of Human Tumor Cell Lines to Sirtinol

Betül Baykal

A Thesis Submitted in Fulfillment of the Requirements

Of the Faculty of Life Science of the Open University (UK)

For the Degree of Doctor of Philosophy

International Centre for Genetic Engineering and Biotechnology

Trieste, Italy

Director of Studies: Dr. Michael P. Myers

External Supervisor: Dr. Vittorio Venturi

September, 2015

DATE of SUBMISSION : 21 SEPTEMBER 2015

DATE of AWARD : 5 NOVEMBER 2015

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEDICATION

I dedicate my thesis work to my brother Ahmet Tarık Baykal, whom I have immense gratitude for being the greatest brother in the world. I thank him for always being there with great love and invaluable support.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. Michael P. Myers, for his excellent guidance, kindness, patience, and for giving me an opportunity for doing research in his lab. I also would like to thank him for encouraging me to make my own decisions and to overcome problems myself during the course of my research.

Thanks to all my great friends and family for their continuous support and encouragement through these years.

ABBREVIATIONS

All the abbreviations can also be found in the text when they introduced the first time.

The abbreviations used only once are not included in this list.

1DE	1 Dimensional Electrophoresis
2DE	2 Dimensional Electrophoresis
ABC	Ammonium Bicarbonate
ACN	Acetonitrile
BSA	Bovine Serum Albumin
CID	Collision-induced Dissociation
DDR	DNA Damage Response
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribo Nucleic Acid
DSB	Double Strand Break
EIF2α	Eukaryotic Translation Initiation Factor 2 alpha
ERLIC	Electrostatic Repulsion Hydrophilic Interaction Chromatography
ESI	Electrospray Ionization
ETD	Electron Transport Dissociation
FA	Formic Acid
FASN	Fatty Acid Synthase
FBS	Fetal Bovine Serum
FDR	False Discovery Rate

GPM	Global Proteome Machine
HDACs	Histone Deacetylases
HDAi	Histone Deacetylase Inhibitors
HPLC	High Performance Liquid Chromatography
ICAT	Isotope Coded Affinity Tags
iTRAQ	Isobaric Tag for Relative and Absolute Quantification
LC	Liquid Chromatography
MALDI	Matrix-Assisted Laser Desorption/Ionization
MEFs	Mouse Embryonic Fibroblasts
MS	Mass Spectrometry
PBS	Phosphate Buffered Saline
PCNA	Proliferating Cell Nuclear Antigen
PPM	Parts Per Million
PPP2R3A	Protein Phosphatase 2 Regulatory Subunit B Isoform R3
PTM	Post Translational Modification
Rb	Retinoblastoma Protein
RP	Reverse Phase
RPC	Reverse Phase Chromatography
SAHF	Senescence-Associated Heterochromatin Foci
SASP	Senescence-Associated Secretory Phenotype
SAX	Strong Anion Exchange
SA-β-gal	Senescence-Associated Beta-galactosidase
SCX	Strong Cation Exchange
SDFs	Senescence-associated DNA-damage Foci
SDS	Sodium Dodecyl Sulfate

SDS-PAGE	Sodium Dodecyl Sulfate, Polyacrylamide Gel Electrophoresis
SILAC	Stable Isotope Labeling with Amino Acids in Cell Culture
TCA	Trichloroacetic Acid
TEAB	Tetraethyl Ammonium Bromide
UV	Ultra Violet

ABSTRACT

My thesis develops and evaluates a novel, hybrid method for mass spectrometry based proteomics, with the goal of increasing the number of protein identifications, while also allowing quantitative analysis to be performed. The hybrid approach I developed uses classical SILAC labeling coupled to a "novel" scheme for enriching cysteic acid containing peptides which are generated by oxidizing the cysteine and cystine residues of proteins. The resultant acidic side chains are used for the selective enrichment of these peptides. As a result, I get the ease of labeling and enhanced quantitative performance of SILAC protocols with the added benefit of the reduction in sample complexity seen with ICAT protocols.

The hybrid method enabled me to study the proteomic profiles of H1229 cancer cells following sirtinol treatment. Sirtinol, which targets members of the sirtuin family, has been proposed as a treatment for neurodegeneration and cancer. In fact, sirtinol induces senescence in H1299 cells, and I was able to observe the long term therapeutic effects of sirtinol at the proteome level.

The optimized hybrid protocol yielded approximately 5 times the proteins identifications from H1299 cells, when compared to the standard protocol. Quantification of this data revealed that 140 proteins are downregulated and 88 proteins are upregulated following sirtinol treatment of H1299 cells. Moreover, an additional benefit of the protocol was an increase in the total number of post translational modifications (PTMs) identified using the hybrid approach. This work will give a better understanding of the mode of action of these compounds, as well as revealing the biological mechanisms of their effects.

In conclusion, the hybrid method is useful for increasing the number of protein identifications from complex samples. It is an accurate and straightforward method for comparative studies. This method is especially useful for drug analysis studies or other binary comparisons.

1. INTRODUCTION

1.1. Overview

Proteomics is a powerful method for biological discovery, because it provides information directly about the proteins in a sample. Proteomics has gained significant importance for understanding disease mechanisms and for biomarker discovery. Together with genomic approaches, proteomic studies hold great promise in identifying the mechanisms of action of new drugs, especially for those drugs that have an indirect mode of action, such as transcriptional regulators. However, it needs to be recognized that proteomics is still an emerging technology that has its own drawbacks [1].

Proteomics is a wide field of study that depends upon effective sample preparation techniques and sensitive instrumentation for the identification of proteins in complex biological samples. Unlike genomics studies, there are no proteomic methods that can identify and quantify the complete set of proteins from a complex biological sample. The general workflow includes several steps, including cell lysis, protein digestion and peptide purification. Given the complexity of the proteome [2], extensive sample manipulation is frequently required before the analytic technique can be applied. In this case the sample complexity is a mix of both the number of distinct components and the variability in their concentrations. There can easily be a million fold difference between the most abundant and least abundant constituents in a complex biological mixture, such as a cell lysate or serum sample.

Although, these issues may be daunting, the aim of proteomics has always been to analyze all the proteins from a sample and this aim continues to be the major force in the development of new instrumentation and techniques. Many different techniques are

available for proteomics studies, however they currently do not effectively meet the need. This creates the necessity for creating new, more effective techniques.

With this perspective, I aimed to develop a novel method for sample preparation that was designed to produce an increase in the number of protein identifications from a complex mixture. In addition, I sought to reduce the number of experimental steps, all while still permitting the quantification of the proteins in the sample.

Secondly, I used this method in comparative studies to investigate the global proteome response to, sirtinol, which targets sirtuins (suppressor of Sirt1). Sirtuins are involved in aging, senescence, neurodegeneration and cancer [3]. Studying this compound is quite interesting, because sirtinol has been investigated as a potential treatment for cancer. Investigating such drugs is important for preclinical studies, because finding the protein targets of a compound is one if the keys to understanding its mode of action.

In this introductory chapter, the main goals and challenges of proteomics will be addressed. The classic proteomics workflows for identification of proteins with mass spectrometry will be explained. An overview of methods employed for comparative and quantitative proteomics analysis will be presented, highlighting the two well-known techniques that the idea of novel hybrid method is derived from. The main idea of hybrid method is presented next, but the detailed steps of optimization and development left for results section.

The second half of my thesis involves the application of the hybrid method for comparative studies to investigate the global proteome changes in biological systems. The effects of sirtinol treatment at the proteome level have been investigated in a cancer model; therefore I have also included an introduction to sirtinol, cancer, and senescence.

1.2. Proteomics

The genetic material of an organism contains the complete set of genes required to build RNAs and proteins. The goal of genomics is to determine and understand the structure and function of a complete. DNA sequence of an organism [4] and this information is often critical for understanding many biological problems, but genomic studies have their own limitations [5]. For example, genomics can only give an indirect measure of protein function, as the levels of gene transcription provides only a rough estimate of protein abundance [6]. Furthermore, genomics does not take into account post-translational modifications that profoundly affect protein activities. Proteomic studies can overcome these limitations since they are focusing on the molecules which have the major functional role in living organisms, proteins [4]. However, the analysis of an organism at the protein level introduces a variety of analytical challenges, many of which still need to be overcome.

1.3. The Challenges of Proteomics

The proteome is the complete set of proteins produced by the genome [7] which varies with time and depends on a variety of physiological or pathological conditions. A proteome contains a wide diversity of unique proteins. In this respect, a proteome is much more complex than genome or the transcriptome that gives rise to it, since proteins are quite dynamic in cell [8]. From the same genome, different proteins are expressed and each protein can be chemically modified in different ways after synthesis, such as carbonylation, phosphorylation and acetylation etc [9 10].

Unlike genomic studies, it is not possible to identify all the proteins in cell with simple methods. Unfortunately, there is no protein equivalent of PCR. Even when high resolution mass spectrometers are used, the sample complexity far exceeds the analytical capacity which results in undersampling of the data. Moreover, the nature of the protein can also

add difficulties to analysis. For example, some proteins, such as highly hydrophobic membrane proteins are more difficult to manipulate [11]. These characteristics of proteins create challenges in proteomics.

In order to deal with these problems, different sample preparation methods or data analysis methods need to be improved. Strategies aimed at reducing sample complexity at the mass spectrometer, or normalizing the dynamic range, are currently being explored to improve the effectiveness of MS based proteomics.

Although, much of the current development of proteomics has focused solely on protein identification, it is clear that quantitative proteomics is particularly important for understanding biological process, as very few biological processes are regulated in an all-or-none fashion. Quantitative proteomics has additional challenges and typically quantitative approaches result in fewer protein identifications then strategies which do not produce quantitative data.

1.4. The Goals of Proteomics

Since the proteins are the actual functional molecules in the cell, studying the proteome is quite important for understanding how a cell functions and how these functions are changing under different circumstances, and proteomics has provided key insights into various types of areas including aging, cancer, drug discovery and infectious disease mechanisms [4].

Proteomics includes the systematic study of the amounts, modifications, functions, localizations and interactions of all or a subset of the proteins found in particular cells, tissues or whole organisms. On a practical level, one of the major goals of proteomics is to provide useful information on mechanisms of human diseases (e.g., cancer) with a long

term aim of developing methods for diseases diagnosis and treatment. However, in a simpler explanation, the main goal of the proteomics is, identifying and characterizing all the proteins in cell and identifying the changes in the proteome in response to changes in cellular state.

Many different methods have been developed to expand the limits of proteomic approaches [1]. The methods of choice can be tailored for the specific goals of the study, such as whether quantification of the proteins is required. Although these methods should be straightforward and easy to perform, they typically require further optimization, and may require additional steps to adapt the protocol to particularly challenging samples and available analytical tools. For example, something as common as blood sera requires additional steps that typically are not required when analyzing whole cell lysates.

Although there are many different techniques, proteomics approaches can be broken down into several common steps such as sample preparation, separation of peptides/proteins, data collection and data analyses. In the next section, these steps and their importance will be explained in greater detail.

1.5. The Classic Proteomics Workflow for Protein Identification

The proteomics workflow involves the experimental techniques and instruments used for acquiring and analyzing the data of a particular proteomics analysis. There are different approaches in proteomic studies. Two main strategies have been used over the years: the 'top-down' and the 'bottom-up'. The top-down strategy is based on the MS analysis of whole proteins [12]. Top-down strategies are frequently targeted at a small number of proteins and are used for the thorough characterization of isoforms, PTMs, etc. of a handful of proteins at a time. On the other hand, the bottom up strategy, which is the strategy used in this thesis, is more frequently used, especially for the analysis of complex

samples. In this strategy the MS analysis is performed with peptides derived from the enzymatic digestion of proteins and protein identification is based on the mapping of these peptide identifications back to their source proteins. Typically bottom-up analysis is much more efficient than top-down approaches and results in many more protein identifications.

Many different workflows can be used for bottom-up proteomic analysis and they have different advantages and disadvantages (Table 1.1.). The majority of these workflows involve, sample preparation, protein separation, data collection and data analysis with mass spectrometry for protein identification **(Figure 1.1.)**. These steps will be covered in next sections highlighting the importance of sample preparation steps, since proteomics studies are highly depending on effective sample preparation and also for giving a better understanding to the objectives of the thesis.

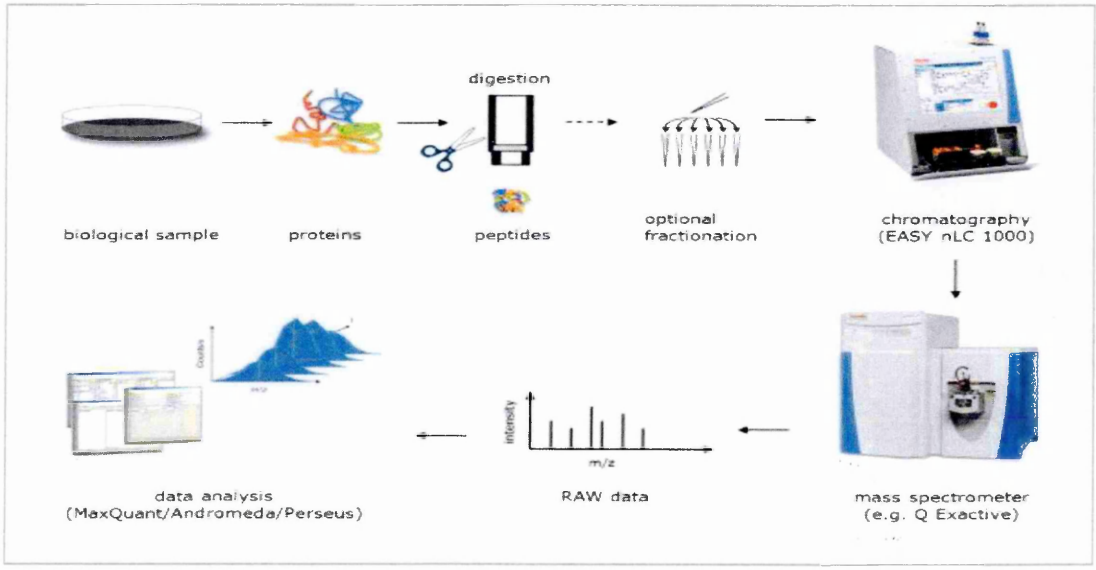


Figure 1.1. General workflow for proteomic analysis. The proteomics workflow entails the isolation of proteins from the biological source, processing the proteins into peptides, one or more separation steps and finally analysis by MS [253].

		PROS	CONS
DIGESTION TECHNIQUES	**In-solution Digestion**	*Easy and fast procedure *Provides good coverage for peptides	*No sample simplification *Frequently needs to be combined with sample cleaning techniques
	In-gel Digestion	*Simplifies complexity. *Orthogonal separation technique to all forms of chromatography. *Efficient contaminant removal	*More time consuming as there are many processing steps *Not all peptides can be recovered
QUANTIFICATION TECHNIQUES	**SILAC**	*Accurate quantification *Minimizes sample to sample variability *True in-vivo technique	*Increases sample complexity
	ICAT	*Decreases sample complexity *Increases protein identification	*Long sample preparation time *Sample to sample variability
	Label-free techniques	*Can be used for many types of samples *Reagents are not expensive	*Sample to sample variability

Table1.1. Pros and cons of some of the proteomic techniques.

1.5.1. Sample Preparation

In proteomic studies, samples can be obtained from virtually any biological source, even from fossilized dinosaurs. Ideally, the sample preparation steps are designed to change

the sample into a form that is suitable for analysis, as biological samples are typically not amenable to direct analysis by mass spectrometry or other analytical techniques.

Sample preparation is the most important step in order to obtain reliable and reproducible results. In general, these steps solubilize the proteins in a sample, while eliminating any interfering compounds such as salts, lipids, nucleic acids etc. There is no universal protocol for sample preparation. There are typically many methods that can be used, or adapted from a known method to fulfill the experimental aims and the methods can be varied depending on the biological sample and the objectives of the study [13].

Sample preparation workflows mainly depend on the nature of the sample such as, type of sample, available amount, hydrophobicity, pH, and temperature stability [13]. The other important factor is the biological question to be addressed. For example, analysis of membrane and/or soluble proteins, post-translational modifications and protein-protein interactions can be studied, but these studies require different sample preparation methods. Similarly, a difficult to extract sample, such as a fossil or the woody pulp of a tree, would require different steps from serum, plasma, or other samples rich in soluble proteins.

1.5.1.1. Sample Solubilisation

In order to analyze the proteome, as many proteins as possible should be extracted from the sample. Successful protein expression analyses via mass spectrometry relies on the solubilisation of proteins from complex mixtures. This can be done by several methods which use combinations of chemical and mechanical disruption strategies. Most common methods involve a chemical lysis using detergents or chaotropic agents coupled with mechanical forces typically generated by sonication or grinding that physically breaks the sample down [13].

To maintain the solubility and stability of the protein samples, many different buffers, reductants, detergents and inhibitors are typically used in the early steps of sample preparation. The classic lysis buffer combination may include detergents like NP40 and Triton-X100 or chaotropes such as urea. Lysis, extraction and denaturation of proteins can be obtained in same step with certain procedures such as boiling and agitating the sample with SDS. It is important to mention that, the choice of lysis buffer combination and mechanical disruption mainly depends on the target proteins and characteristics of the sample source. In addition, proteases might be activated during protein solubilisation. This can be avoided by adding protease inhibitors or by keeping all the solutions ice cold. There is no the absolute best way to lyse a sample but speed and cleanliness of preparation definitely affect protein identification.

1.5.1.2. Removal of Contaminants

Once proteins are extracted from their source, many contaminants need to be removed because most of them are not compatible with the downstream steps, especially the LC-MS. For example, many extraction reagents, such as strong detergents will interfere with enzymatic digestion and most detergents interfere with reverse-phase separation and mass spectrometry [14]. The removal of other contaminants such as, salts, nucleic acids, lipids, polysaccharides is also essential for MS analysis since these contaminants cause chromatographic interference and signal suppression.

Removal of contaminants such as DNA, lipids, detergents etc. are most often removed via dialysis, ultrafiltration, precipitation with TCA or organic solvents and solid-phase extraction techniques and have been incorporated into commercially available clean-up kits [13].

One common approach for contaminant removal is precipitation. There are various techniques for precipitation including using strong acids or organic solvents. Protein precipitation frequently causes selective sample losses, especially as solubilizing some proteins can be challenging. Similar selective sample losses are seen with other removal techniques, such as ultrafiltration. Therefore, different contaminant removal techniques are continually being introduced and are often tailored towards specific analytical challenges.

1.5.1.3. Digestion Strategies

Protein digestion can be performed using in-gel, in-solution, in-filter approaches, etc. In-solution and in-gel digestion are the two most widely used techniques to prepare bottom-up proteomic samples **(Figure 1.2.)**.

In-solution digestion is a straightforward, yet extremely powerful technique that typically used in shotgun proteomics. It involves, denaturing, reducing, alkylating and digesting the proteins in liquid phase. Generally, in-solution digests are fractioned after digestion but fractionation can be also performed before digestion using different forms of chromatography (i.e., ion exchange, reverse-phase). Since the sample typically does not require a change from liquid phase, it is easier to minimize sample losses during the digestion phase.

SDS-PAGE is probably the most widely used technique today, since it has the ability to separate thousands of proteins in a single sample from a complex mixture [15]. In-gel proteomics, involves separation of the proteins before digestion, using acrylamide gels and then processing the protein containing bands for analysis by mass spectrometry. In addition to fractionating the sample, SDS-PAGE is also incredibly useful for removing unwanted contaminants. It is also relatively simple procedure and the gel also acts as a

good storage for proteins for future analysis [16]. Gel based sample preparation is almost always used for bottom-up proteomics, because the full proteins are difficult to elute from the gel, while the much smaller peptides produced by digestion are much easier to recover from the gel [17].

Many different endoproteases can be chosen for protein digestion but trypsin is widely used, because it generates peptides that are highly compatible with MS analysis. Trypsin cleaves C-terminal to arginine and lysine [18]. This typically generates peptides in the 800 Da to 2000 Da range, which can be analyzed with high sensitivity via MS. Moreover, because of the basic nature of lysine and arginine, tryptic peptides are efficiently ionized during mass spectrometry, due to the presence of a basic residue at both the N- and C-termini of the tryptic peptides.

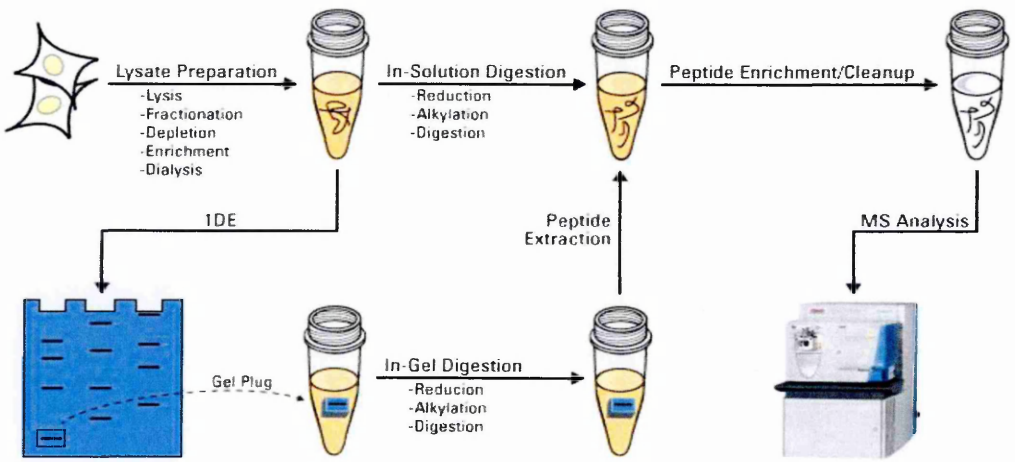


Figure 1.2. A comparison of in-solution and in-gel digestion strategies in proteomics. Gel electrophoresis is used to separate whole proteins. Bands containing proteins are excised from the whole gel and subjected to enzymatic digestion and peptide extraction. However, in the in-solution protocol, the proteins are digested directly from the solution containing them. In both cases, the resulting peptides are analyzed with mass spectrometry [19].

1.5.2. Fractionation and Separation of Proteomic Mixtures

Fractionation and separation techniques for proteomic mixtures are used to separate the components of a mixture in order to make the identification of these components more efficient. As mass spectrometry is a time dependent analysis technique, these extra steps of separation are used to provide more time for the MS to analyze a single sample. In this regard, the separation is done in order to reduce the complexity of a sample entering the MS, so that any given time, the MS is not dealing with all the components at once.

Different characteristics of the components can be used for separation of mixtures such as polarity, electrical charge and molecular size. In mass spectrometry based proteomics, separation is usually performed with liquid-chromatography (LC). This on-line separation, helps reduce the complexity of a sample entering the MS, as the components of each "fraction" are dealt with independently from other resolved fractions. This separation is necessary, because even the digest of a single protein is too complex to allow the MS to analyze all the resulting peptides. Therefore, sample separation is one of the crucial steps for MS-based proteomics [13].

Complete separation of highly complex mixtures is technically impossible, as there will be some overlap of components with similar characteristics. The main goal is to provide enough separation to allow the MS ample time to analyze the co-eluting components. A single chromatographic step is typically not sufficient for separating digests from more than a handful of proteins at a time. This is especially true, when some components, PTMs for example, are present at a much lower stoichiometry. Multidimensional chromatography techniques can be used for these situations. In this case, the peptides are fractionated by two or more steps of chromatography that are designed to have the least amount of co-eluting components as possible. Ion exchange chromatography is often

coupled to reverse phase chromatography as these typically have completely different chromatographic profiles. In fact, most gel based workflows can be considered multi-dimensional. The need for fractionation and separation is a result of the complexity of the samples.

1.5.2.1. Sample Complexity

Sample complexity refers to two main characteristics: the diversity of proteins in the sample and the range of concentration of these proteins. A typical proteome can contain thousands of unique proteins. Coded from 20.000-25000 genes [20], it is estimated that human body can contain over a million unique protein products [21] that result from alternative splicing, proteolytic processing, etc. This number goes even higher if the various types of PTMs are considered. For example, human plasma has been estimated to have more than 10.000 core proteins [22] and only small fraction of it can be efficiently characterized with current technologies. This is not just an effect of the complexity in number of proteins but is also impacted by the large disparity in their concentrations.

Proteins from biological samples typically display a large dynamic range between low and high abundance **(Figure 1.3.)**. The range of abundances of these proteins can be quite high [2]. Plasma proteins can have 10^{12}-fold concentration range, from millimolar (for albumin) to attomolar or lower for certain cytokines [23]. This large dynamic range swamps the linear range of current proteomics methods, which is typically between 3 and 4 orders of magnitude [24]. Selective enrichment or depletion of components has been used to help alleviate the problem of the naturally occurring dynamic range, but even a 99.9% depletion of albumin, results in albumin still being one of the most abundant components of plasma.

When it comes to mass spectrometry, the sensitivity for detection and identification of peptides is strongly affected by sample complexity. For instance, the highly abundant peptides will be identified much more easily and they will shadow the less abundant ones. This is largely due to the fact that in most MS systems, there is a competition for ionization and highly abundant peptides suppress the signal from less abundant ones. This is a general issue with all sorts of proteomic methods and increased separation combats this problem by decreasing the overlap of between high abundant components and other components in the sample.

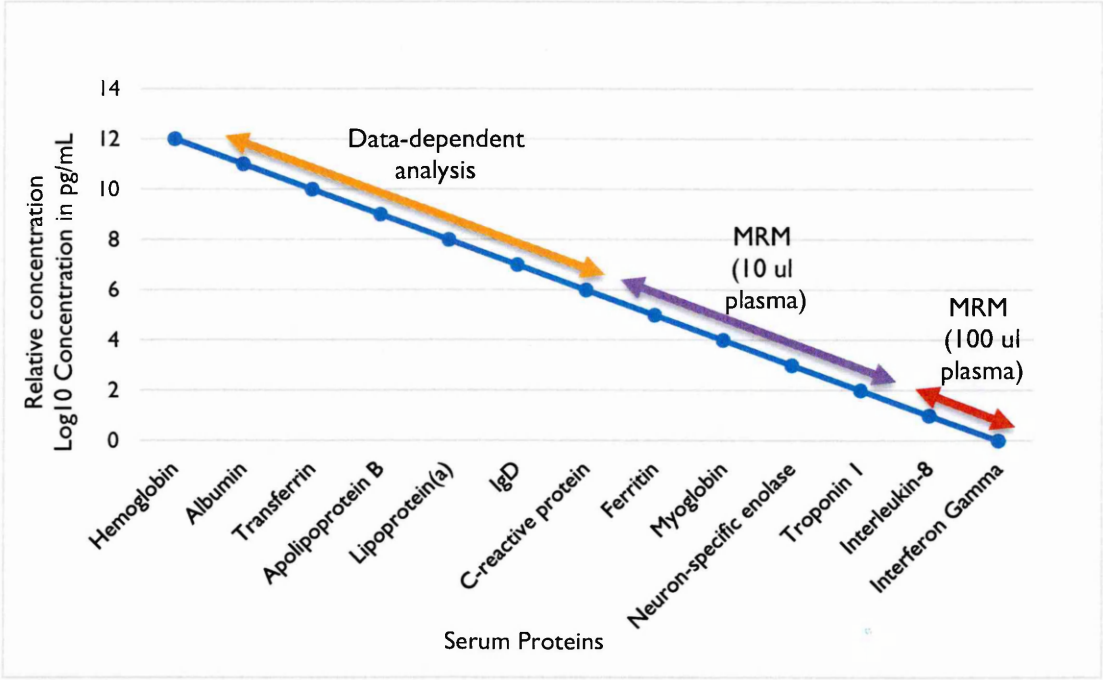


Figure 1.3. Proteins in serum have a large dynamic range. The concentration of various serum components has been estimated using mass spectrometry, Multiple Reaction Monitoring (MRM), and other analytical methods. Adapted from "Proteomic retrenches" by P. Mitchell, 2010 Nature Biotechnology 28, 665–670 [2].

1.5.2.2. On-line Separation of Proteomic Mixtures

Many different mass spectrometry and pre-fractionation combinations have been developed to improve the analysis of complex proteomic samples. High-performance liquid chromatography (HPLC) is a classic separation technique which provides a wide

variety of separation modes. This has been improved with ultra-performance liquid chromatography (UPLC), which provides an increased resolving power and speed of separation [25]. HPLC and UPLC use the same functional principles for separation, with UPLC having highly optimized columns and flow paths. The optimization is almost solely aimed at reducing column particle sizes and reducing the extra column dead volumes, which result in "ultra-high" pressure systems.

In bottom-up proteomic analyses, the most used HPLC/UPLC stationary phase is the reverse-phase column [26]. It is usually used as the last separation step before MS because the chromatographic buffers are highly volatile and are ideally suited for interfacing with a MS instrument. RP-HPLC separates peptides according to their partition coefficient between a polar mobile phase and a hydrophobic stationary phase. In this type of separation, proteins and peptides are bound to the stationary phase and the separation carried out by application of gradually increasing concentrations of an MS compatible organic solvent. In addition to using MS compatible solvents, RP-HPLC is one of the highest capacity and resolution chromatography modes for peptide separation.

While reverse-phase columns are widely-used, many other stationary phases can be used for proteomic analysis in one and two-dimensional separations. A separation technique, known as electrostatic repulsion and hydrophilic interaction chromatography [27] (ERLIC) is a recent development, and has been used specifically for phosphoproteomic studies. ERLIC uses two characteristics for separation: hydrophilicity and charge state. ERLIC usually uses low-organic and high–pH gradient as the initial mobile phase and moves to high organic and low-pH mobile phase to separate the bound peptides. Hence, ERLIC elutes peptides in order of increasing hydrophobicity and acidity. Other separation modes include, ion-exchange chromatography, exclusion chromatography [28]. The separation can be performed with column effluent flowing directly into the instrument (online

separation) or fractions can be collected prior to analysis on the instrument (offline separation). Online separations tend to be faster and have less sample loss, while offline separation allows for modes that are not directly compatible with MS to be incorporated into MS based workflows and many multiple dimension LC-MS/MS protocols are based on performing the first separation offline and the second separation online.

Multiple dimension separations greatly reduce the sample complexity entering the MS, especially when the two dimension are orthogonal. The number of peptide and protein identification is greatly increased when two or more chromatographic systems based on different selectivity are used together before MS analysis (2D-LC) [29]. Strong Cation Exchange (SCX) chromatography coupled to Reverse Phase Chromatography (RPC) is one of the most popular pairings [30]. In this system, the peptides are first separated based solely on their charge state in the first dimension and then these fractions are separated based on their hydrophobicity by RP-HPLC. Since the charge state and the overall hydrophobicity are independent, only a small subset of the peptides co-elute on both columns [30]. This type of chromatographic system allows for a high number of identifications since the resolution (the power of separation) increases because of two dimensions of separation. However, this is also coupled to an increase in the time it takes to run a single biological sample, and this increased time is also critical to allow the MS increased time to analyze each component.

1.5.2.3. Reducing Sample Complexity by Off-line-Fractionation of Proteins

Reducing sample complexity can be accomplished as part of the sample preparation steps using, immunoprecipication, immunodepletion and fractionation of proteins with SDS-PAGE. Some of these methods, such as immunoprecipitation, are targeted reductions [31] that enrich a target protein and its interactors at the expense of proteins that do not

interact with the affinity matrix. Similarly, removal of the most abundant proteins can be performed to reduce protein concentration variability. For example, albumin can be removed from samples by immunodepletion [32]. This allows proteomics analysis to begin to probe proteins that are initially present at much lower concentrations. However, potentially significant proteins can be lost from the final sample due to non-specific binding or other factors [33]. Non-targeted approaches include methods where the "whole" sample ends up on the MS. For example, if an entire lane from a gel is analyzed, chromatographic prefractionation where every fraction is analyzed [34], and affinity approaches such as Isotope Coded Affinity Tag (ICAT) [35] where specific peptides are enriched without regard to their protein of origin.

An often overlooked form of untargeted protein separation is gel electrophoresis (e.g., SDS-PAGE) before MS analysis[36]. SDS-PAGE is still the highest resolution technique for proteins [37]. Since this separation is based on whole protein size, it is orthogonal to any peptide based separation performed on the digested gel pieces. After separation with on-line or off-line techniques data is acquired by a mass spectrometer.

1.5.3. Data Acquisition in Mass Spectrometry Based Proteomics

Data acquisition is the third step in proteomics workflows. It includes the ionization of the samples and these ions are then analyzed by the mass spectrometer. Importantly, mass spectrometers can only analyze charged species.

Mass spectrometry is a technique in which molecules are ionized and their mass-to-charge ratio (*m/z*) is measured [38]. The *m/z* ratio is calculated by simply dividing the mass of an ion by the net number of charges it carries. A single protein or peptide may possess a variety of *m/z* ratios, depending the distribution of charge states. It is not uncommon for even short peptides to produce +1, +2, +3, forms.

Identification of peptides and proteins in a mass spectrometer depends on the ability to control the behavior of these molecules. For example, in a typical instrument, electrospray is used to produce the ions, which are accelerated into the mass spectrometer via high voltage. Ions with a specific m/z can be selected by deflecting them with a magnetic field. Each m/z will be deflected differently resulting in a separation of the ions. The selected ions are directed at a detector resulting in the formation of a peak for the selected m/z ratio. These procedures are done in the mass spectrometer's components which are the ion source, the mass analyzer and the detector **(Figure 1.4.)**.

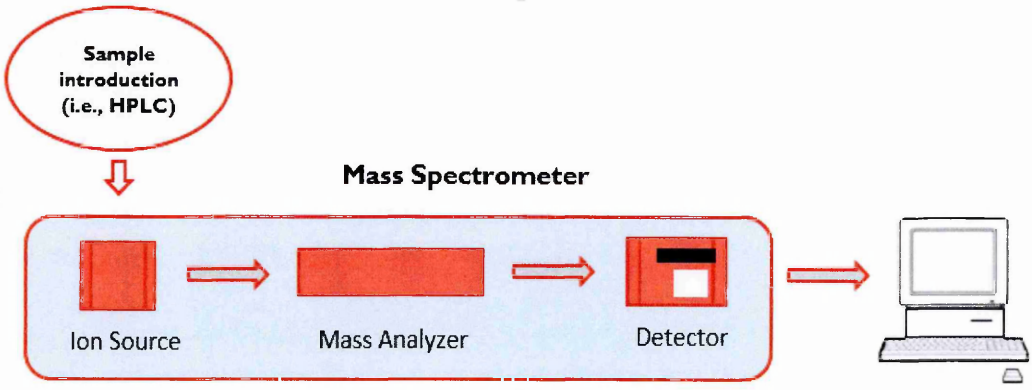


Figure 1.4. The components of a mass spectrometer. A schematic of a typical mass spectrometer, showing the flow of ions from the ion source, through the mass analyzer and finally to the detector.

The ion **source** is an ionization chamber where the stream of ions is generated. There are different ionization methods, including, electron impact ionization (EI), chemical ionization (CI), matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), etc. MALDI and ESI are the predominant modes for biological mass spectrometry since they are considered 'soft' ionization methods that do not destroy fragile biological samples. Since my work solely relied on ESI as an ionization source, the emphasis of the discussion will be on this method.

ESI is an evaporation based ionization mode. It includes the generation of charged liquid droplets as a fine spray. It is a form of atmospheric pressure ionization (API) since the ionization does not take place under vacuum. In this mode, the sample is in a liquid matrix, which is ideally a mixture of aqueous and organic solvents. Typically the sample is sprayed from a very fine needle, which maintained at high voltage (1-5 kV) to generate charged liquid droplets. The high voltage concentrates the charge in the flowstream at the tip of the needle. Electrostatic repulsion at the tip of the needle causes the flowstream to "explode" into fine droplets **(Figure 1.5.)**. A heated gas (e.g., nitrogen at 70 °C) is used to accelerate the evaporation of the solvent, resulting in the transformation of the analyte from liquid phase to gas phase. The residual charge in the droplet is transferred to the analyte and the ions are directed to the mass analyzer by the high voltage.

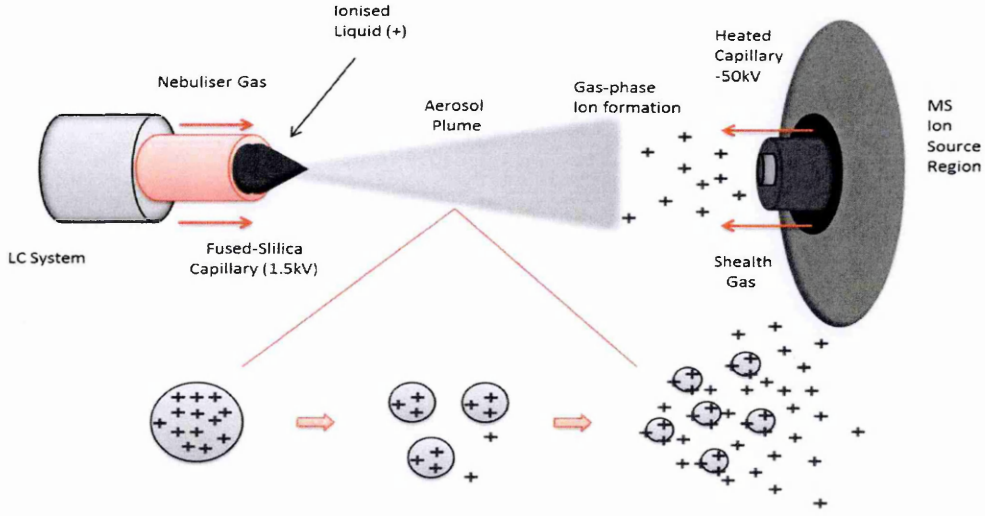


Figure 1.5. Schematic of an electrospray ionization source. A fine spray is formed at the tip of the needle and the plume of ions is drawn into the mass spectrometer by the high voltage field [19].

The **analyzer** is where the m/z ratio of the analyte is measured. Different types of mass spectrometers do this differently and they all use different properties of an ion in an electric field for this measurement. In addition, the analyzer also directs the ions to the detector. Many types of mass analyzers have been used for proteomics including:

Quadrupoles (Q), time of flight analyzers (TOF), ion-traps (IT), and Fourier transform based instruments, such as the Orbitrap™. Each mass analyzer has its own unique set of advantages and drawbacks [39]. For example, ion trap analyzers consist of three electrodes: two end cap electrodes and one ring electrode **(Figure 1.6.)**. An oscillating voltage is applied between the ring electrodes and as ions pass into this analyzer, they are trapped by the electrostatic field created by these electrodes. When the amplitude of the field within the device is changed, certain trapped ions begin to resonate within the trap, ultimately resulting in them leaving the trap where they can be detected. Each *m/z* has its own characteristic resonance amplitude, so it possible to selectively remove or retain ions from the trap based on their *m/z*. This behavior allows for the production of a mass spectrum, as well as allowing for the isolation of a particular analyte (*m/z* ratio) for further study inside the ion trap.

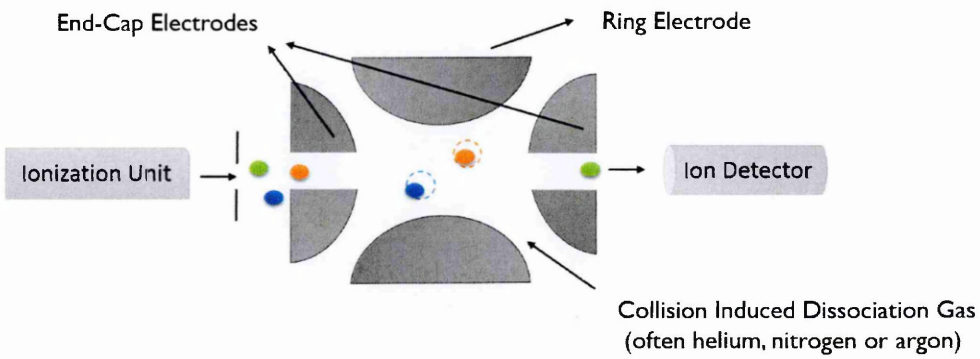


Figure 1.6. General scheme of ion trap analyzers. Ion trap analyzers consist of three electrodes: two end cap electrodes and one ring electrode and ions produced at the ion source are retained in the trap where they can then be manipulated.

For example, the analyte can be fragmented within the mass analyzer and then the fragment masses can be determined. This is typically referred to as MS/MS mode. MS/MS is the favored mode of analysis for the identification of peptides as the fragments have a higher information content then is contained in the m/z ratio of the originating, or

precursor ion, and there are a variety of methods to determine the sequence of a peptide from the fragment ions it produces in a mass analyzer.

It is often assumed that in order to identify each peptide, the sequence of the peptide needs to be obtained. To be able to do this, mass spectrometers typically make two separate mass measures. The first to determine the components currently eluting from the column, and the second mass measure to determine the fragmentation pattern of one of the components. In the second run, the peptide ions are typically resonated or otherwise collide with molecules of an inert gas, in a process called Collision Induced Dissociation (CID).

How peptides break into pieces, their dissociation pathways, mainly depends on the actual peptide sequence. However, depending on the fragmentation method used, different fragment ion types can be produced. When peptides break apart during CID, the charged ions are retained on one or more of the fragments. These charged fragment ions are measured by the mass analyzer, and this forms the fragment mass spectrum. In an ion trap mass analyzer, peptides tend to fragment along the peptide backbone, which produces a series of ladder ions that reveal the peptide's amino-acid sequence. When the N-terminal fragment retains a proton, the fragment ions are typically named 'a' or 'b' ions. The other predominantly formed fragment ion, the y ion, forms when the C-terminal fragments retains a proton. CID mostly form a-, b- and y-ions **(Figure 1.7.)**, while other modes of fragmentation produce other classes of ions, for example chemically induced fragmentation by Electron Transfer Dissociation (ETD) produces mostly c and z-ions. Fragments with a, b, or c in the name denote ions with a proton on the N-Terminus while those with x, y, or z are the corresponding ions with the proton at the C-Terminus.

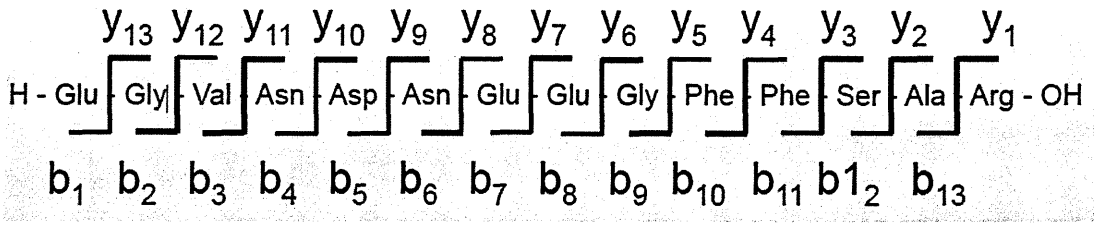


Figure 1.7. Theoretical backbone fragmentations for Glu-Fib. b and y ion series is shown using the nomenclature of Roepstorff [254].

The detector is the third component of a mass spectrometer. On ejection from the ion trap, ions are directed to **the detector.** The ions strike the detector which amplifies this signal. The strength of that signal is directly related to the number of ions that strike the detector. In the case of an ion trap mass spectrometer, this results in an ion count for each displaced *m/z*.

A mass spectrum is a histogram of these detector based ion-counts where the Y-axis represents the number of ions impacting the detector and the z axis is the *m/z* ratio **(Figure 1.8.)**. In other words, it is a record of mass-to-charge ratios (*m/z* values) and intensities of all the ions which were trapped or generated in the ion trap [40]. In a specific proteomic experiment, this mass spectrum represents the ions that were trapped at the time they were scanned out to the detector.

After the acquisition, the spectra are analyzed with 'peak detection algorithms', to simplify the spectra by removing noise, removing multiple charge states for the same compound and deisotoping finds the mass of the peptide that does not contain heavy isotopes of carbon, nitrogen, oxygen etc. Eventually a peak lists is created for further analysis. In these lists, each peak is typically denoted by its m/z, charge state, and intensity. These peak lists are then used to identify the peptides found in the sample.

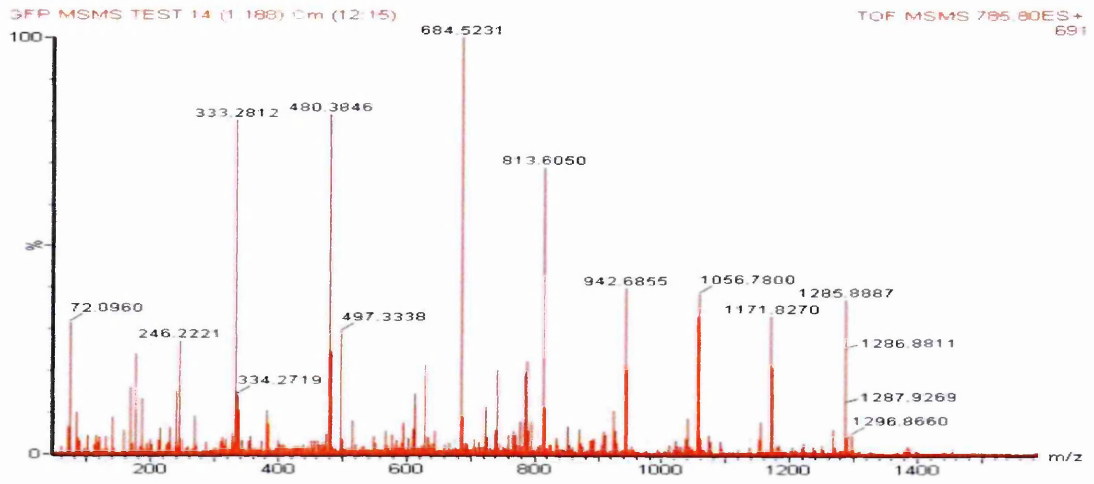


Figure 1.8. A representative mass spectrum. MS/MS spectrum of Glu-Fibrinopeptide (EGVNDNEEGFFSAR).

1.5.4. Data Analysis in Mass Spectrometry Based Proteomics

In a typical proteomic experiment, the mass spectrometer can generate thousands of MS/MS spectra per hour and manual interpretation is not a practical option [41]. Several computational approaches and software tools have been developed for automated assignment of peptide sequences to MS/MS spectra. There are roughly three different methods to make this match. In the first method, peptide sequences are directly extracted from spectra, without referring to a sequence database for help [42]. This is called as, '*de novo*' sequencing. This method allows identification of peptides whose exact sequence is not in the database, but it is computationally intensive. As CID produces a series of ladder fragments, the sequence can be determined by matching the differences between peaks to the masses of the amino acids. This is much more complicated than it sounds, as there can be hundreds of ions in a spectrum compounded by 20 amino acids, compounded by 3 common types of ions that can form, compounded by spurious peaks, etc. However, if a fairly complete sequence database is available for the source organism, then other more efficient methods can be used. For example, peptide identification can be obtained by

correlating experimental MS/MS spectra with a sequence database. In this case, the spectra are turned into a series of short amino acid sequences, and then compared to a sequence database. This approach allows for a less sophisticated de novo interpretation of the spectra to be performed as you are relying on the sequence database to find the best interpretation [43]. The most efficient and popular search engines are based on spectral matching, where the sequence database is converted into theoretical spectra, which are then compared to the experimentally derived spectra [44]. The main differences between search engines is how they score the matching and how they separate statistically significant matches from random matches.

1.5.4.1. Peptide Identification by MS/MS Database Searching

There are a growing number of MS/MS database search tools that are available for use, such as SEQUEST [43], MASCOT [45], X!Tandem [46], OMSSA [47]. These tools operate in a similar manner. They take an experimental spectrum and compare it against the theoretical fragmentation patterns established for peptides derived from the searched database. A single experimental spectrum may be compared to thousands of theoretical spectra as typically all theoretical spectra with a mass similar to the precursor mass are compared. Therefore, the search does not involve all the database but a subset of it, based on user specified criteria, such as mass tolerance, proteolytic enzyme constraints and types of posttranslational modifications allowed, etc.

Theoretical fragmentation patterns are calculated for each of the candidate peptide using common peptide fragmentation rules. List of matches (peptide sequences) are the output of the database search tools, which are ranked according to the scoring scheme used by each particular tool. The best scoring peptide match is assumed to have the highest possibility if being correct.

Several parameters are commonly used with MS/MS database searches. These would include; types of fragment ions to include (such as c, b, y, z-ions), peptide ion charge state, parent ion mass tolerance (depends on the mass spectrometer used), enzymatic digestion constrains (i.e., trypsin cleaves after K and R), and chemical or posttranslational modifications. Appropriate choices in search parameters is quite important, since it can affect the quality of the output and also the time required for the database search [38].

1.5.4.1.1. Databases

Many sequence databases are available. They vary in completeness, redundancy and quality of sequence annotation. The choice of which database to use mainly depends on the goals of an experiment. For example, if the identifications of polymorphisms are important, the search should be done against a very complete, highly redundant database. However, overall protein identification may be more efficient using a database with less redundancy, as the extra copies of sequence are difficult for the statistical algorithms to interpret. Therefore, large, redundant databases introduce more false identifications. In fact, if the identification of sequence variants is not crucial, it is better to use a well-curated and well-annotated databases such as UniProt or RefSeq. Often times a well-curated, redundant database, such as the IPI databases offer a good balance between completeness and redundancy.

1.5.4.2. Scoring Schemes and Evaluation of Results

When MS/MS spectra are searched against a sequence database, the best scoring peptide assignment is returned from the database search tool with no guarantee of its correctness. For more accurate analysis, search results typically undergo a statistical analysis that is used to remove false identifications [41].

The raw database search score is a score that indicates the degree of similarity between the experimental spectrum and the theoretical spectrum. However, it is important that these raw scores be given a statistical framework to aid in filtering out false identifications. Since Mascot and X!Tandem is used in this thesis, the emphasis will be on this two search tools.

Mascot [45] computes a probability-based score called an ion score. The Mascot scoring scheme is proprietary, but it has been described in general terms. Instead of reporting the number of matched peaks, Mascot estimates the probability of that number of matches occurring by chance given the number of peaks in the searched spectrum and the cumulative distribution of *m/z* values of predicted ions for all candidate peptides in the database. For example, a matched ion that is found in many theoretical spectra, for example the fragment ion for lysine, does not add as much statistical weight to the score as a match to a much less common ion. Mascot then performs this comparison iteratively with more and more experimental ions, in an attempt to maximize the ion score.

On the other hand, X!Tandem [46], calculates each peptide assignments by a statistical confidence measure called the expectation value. The scoring functions in X!Tandem is based on the shared peak count approach, where only peaks shared between the experimental and theoretical spectra are counted. The process of validating peptide assignments to spectra is assisted by the conversion of the search scores into expectation values [48] . A smaller expectation value means that the match is less to have occurred by random chance. X!Tandem assumes that the highest scoring match is correct and then uses all the other "wrong" matches to calculate the expectation value. In practice an expectation value is similar to the more commonly used p value, but the expectation value also takes into account the number of trials. However, even with statistical based scoring, the best match obtained from database search tool is not necessarily correct and there

are many reasons why database tools can fail to assign correct peptide sequences to MS/MS spectra.

Even when the correct peptide sequence is in the database, another peptide can score higher than the correct one owing to deficiencies of the scoring scheme used. Scoring schemes implemented in the database search tools involve a simplified representation of the peptide ion fragmentation process. Since it is currently impossible to correctly predict the intensity of each fragment ion, the theoretical fragmentation spectrum sets all the fragment ion intensities to be the same. Furthermore in experimental spectra, the intensity of a fragment ion depends on the amino acids located on the sides of the corresponding peptide bond and fragmentation frequently fails to form between some bonds, resulting in drop outs from the experimental spectra and results in incorrect peptide matches [38].

Correct interpretation of MS/MS spectra is also dependent on spectrum quality. When a spectrum contains very few high intensity peaks or has missing fragment ions peaks, owing to poor fragmentation of the peptide, the spectra is considered low quality. Peaks from low quality spectra contain less information about the sequence of the peptide, they are more likely to result in incorrect sequence assignments. Importantly, sample preparation can play a key role in maximizing the number of high quality spectra [41]. Another significant source of errors can occur when assigned peptide sequences are mapped back onto their proteins of origin.

1.5.4.3. Protein Inference

Identifying proteins, can be managed by grouping the peptides according to their corresponding proteins. There are several difficulties in assembling peptides into protein,

principally due to distinct proteins often producing peptides that are also produced by other proteins.

Peptides that are correctly identified tend to group to a smaller number of proteins, as the digestion of most proteins produces many peptides. Furthermore, the number of peptides identified in a protein is correlated to its abundance in the sample and the size of the protein. However, the peptides that are incorrectly assigned typically do not cluster within a small subset of proteins and are distributed more of less randomly across the database [49]. As a result, incorrectly matched peptides cause a high error rate at the protein level [49] and this problem is often referred to as the protein inference problem, as the presence of a protein is "inferred" from the identified peptide sequences. The nonrandom grouping of correctly matched peptides has been used to help solve the protein inference problem [50].

Identification of a shared peptide, whose sequence is present in more than one entry in the protein sequence database, makes it difficult to infer which of these candidate proteins was actually present in the sample. These redundant peptide sequences are mostly due to the presence of homologous proteins, splicing variants, and poor annotation of the database [49,51]. This issue is especially important for organisms where alternative splicing generates many related protein sequences.

The difficulties mentioned above are addressed in the computational tool ProteinProphet[49]. This program uses the nonrandom grouping of correctly assigned peptides to help solve the protein inference problem. Peptides corresponding to single hit proteins are penalized while peptides corresponding to multihit proteins are rewarded. Therefore a protein identified by single peptide will only be included if the spectral match is very high, while a protein with multiple peptide matches will include spectral matches of much lower quality. In addition, the shared peptides are apportioned

among all corresponding proteins, and a minimal protein list of proteins is produced that is sufficient to account for all observed peptides. For example, if there are 10 peptides that map to 2 variants of actin, and α-actin accounts for 9 of these peptides and β-actin includes all 10 peptides, then only β-actin will be included in the final list. The output file includes a list of proteins, their computed probabilities, the matching peptides, and any annotations coming from the searched protein sequence database. ProteinProphet is a powerful way to remove incorrect peptide and protein assignments and has been integrated into a number of other tools, including those used for quantitative analysis such as XPRESS [52] and ProteoIQ [53]. With these programs, relative protein abundance ratios can be obtained in case of quantitative proteomics experiments are performed with ICAT, SILAC or similar techniques. Importantly the quantitative data can also be used to help solve the protein inference problem, as the relative abundance of a peptide can be used to map it back to the correct protein, because all the peptides from that protein should have a similar quantitative value.

Quantification of the proteins in a sample is important, because most biological process are not regulated by the presence or absence of a protein, but based on its specific activity or concentration in the sample.

1.6. Quantification of Proteins

One of the goals of proteomics is the identification of proteins in given sample. While proteomics is often used in these qualitative studies, there is also a growing need to be able to quantitate the protein levels. Quantitative proteomics is often considered more complex than the qualitative identification of proteins, but quantitation is required in order to understand the kinetics and biological mechanisms of most normal and pathological processes.

There are many approaches for protein quantification. For example, ELISAs and radioisotopes, and even western blotting have been used for quantification. Despite the high sensitivity of these approaches, they cannot be used for high throughput studies.

However, MS has become the default method when the quantification of many proteins is required. Importantly, MS can be coupled to many types of quantification schemes. For example, MS can be used to identify proteins, while the quantification can be performed in-gel by measuring the color intensity after fixation of dyes to the protein. In much the same way that radioactive isotopes have traditionally been used for quantitation, mass spectrometry allows for quantitation to be performed by labeling with stable, non-radioactive isotopes [54]. In addition to these labeling approaches, it is also possible to perform MS based quantitation in a label free manner **(Figure 1.9.)**. Each approach has some advantages and disadvantages and these will be discussed in the next sections.

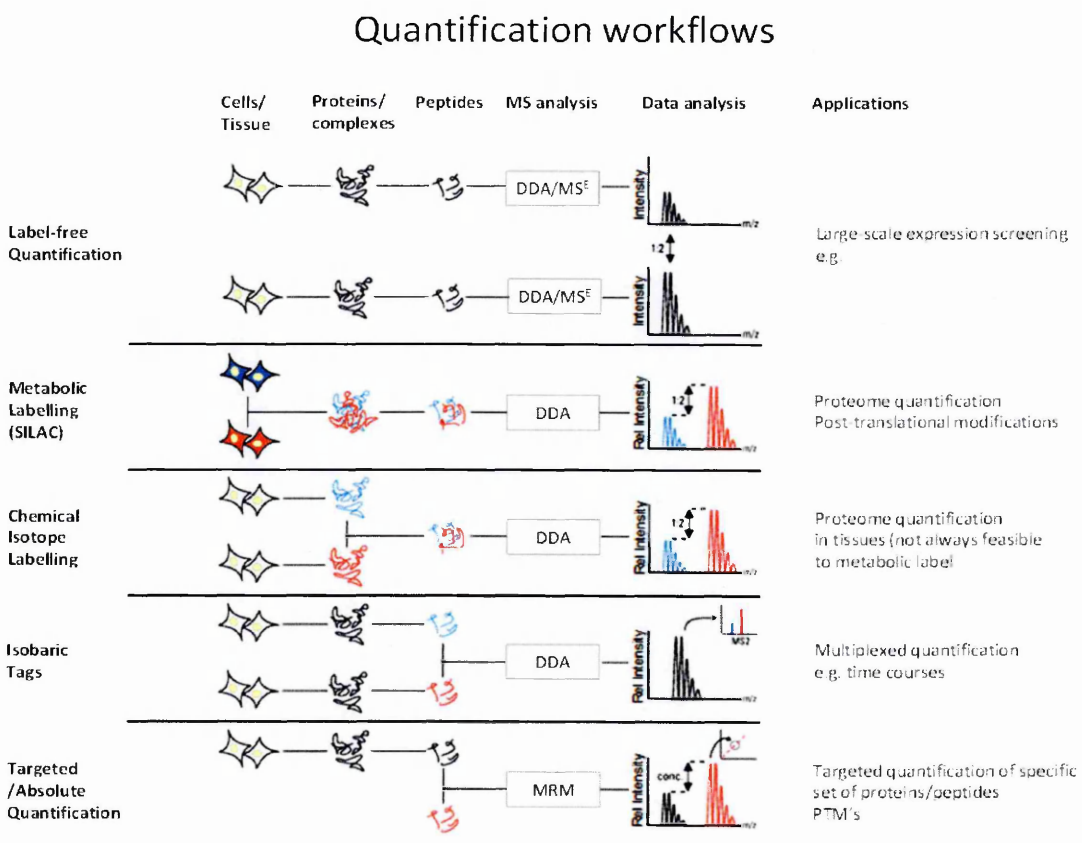


Figure 1.9. Diagram of different quantitative workflows in proteomics [255]. MS can be used to quantitate proteins using label free approaches, such as peak counting, or labeled approaches, such as SILAC, where stable isotopes are used to track how much of each peptide comes from each sample. In addition, targeted approaches such as MRM are used when the proteins to be quantitated are already known. This is typically used in biomarker workflows or when absolute quantification is required. Almost all these techniques use Data Dependent Acquisition (DDA) where the precursor scans are used to determine further analysis.

1.6.1. Gel-based Quantification of Proteins

Gel-based quantification involves electrophoresis [55]. Gel electrophoresis for proteomics involves a porous polyacrylamide medium where proteins migrate according to their physicochemical properties under the influence of electric fields [55]. The proteins can be separated via different properties, for example separation based on size is the most commonly used technique.

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is the most used. The tris-glycine discontinuous buffer system, originally perfected by Laemmli, is still in widespread use [56]. In this technique, SDS binds non-covalently to the proteins and simultaneously denatures them and normalizes their charge state [57]. Therefore, the proteins act like they have identical charge and shape, and their electrophoretic mobility principally depends on their molecular mass.

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) separates proteins by two successive rounds of electrophoresis: the first round separates the proteins based on their isoelectric points, as they stop moving in the gel when they obtain a net neutral charge and the second round uses SDS-PAGE to separate the proteins by their molecular mass. Since 2D separation has a better resolving power than 1D, it has been classically used for the quantification and identification of proteins from complex samples [58].

1.6.1.1. Staining of Proteins for Gel-Based Quantification

The protein spots or bands can be detected after electrophoresis by direct in-gel staining [59]. Commassie brilliant blue and silver nitrate, are two of the most commonly used protein stains, but fluorescence based dyes are also available. Importantly, the fluorescent based dyes can be customized in unique ways, such as for the detection of glycosylation or phosphorylation.

Proteins can also be stained prior to electrophoresis. This system is used in the 'Difference gel electrophoresis' (DIGE) method that needs only one gel to detect and quantitate three samples labeled with three different fluorescent dyes [60]. However, not all fluorescent dyes can be used in this approach, as the dyes can change the mobility of the proteins on the gel. Typically fluorescent stained proteins are detected with digital imaging and then quantitated using specific software tools.

1.6.1.2. Advantages and Limits of In-gel Quantification

Gel based methods typically produce a stable image of the protein pattern, which should be reproducible. Importantly, the proteins are stabilized by being encapsulated in the gel matrix, allowing for both protein diversity and PTMs to be preserved so they are available for further studies, such as MS. One of the principle problems with gel-based approaches is that they typically require many repeats due to the variability between gels. This is also compounded by other problems, such as protein overlap, and the need to process many spots or bands for mass spectrometry, as there is little value in seeing a change in abundance on a gel without a corresponding protein identification. Even when a large number of spots are identified, the total number of proteins identified is much lower than expected, because proteins typically run as multiple spots in 2D gel systems.

1.6.2. Gel-free Quantification of Proteins (Peptide-based Quantification)

Peptide-based quantifications by MS are widely used in proteomics. In this approach, proteins in a sample are digested enzymatically and the resulting mixture of peptides are separated and subsequently analyzed by MS. This approach actually increases the sample complexity, because single proteins are converted into many peptides via enzymatic digestion. To deal with this problem, fractionation of the proteins or peptides is performed. The complexity can also be reduced by subcellular fractionation [37].

Once the sample has been appropriately prepared, peptide-based quantification of proteins can be carried out by MS. Two approaches are generally used: label-free quantification or labeled quantification. A variety of quantitative information can be used for label-free quantification. For example, the relative abundance of a peptide can be determined by comparing the peak area during LC-MS. For example, the SILAC data

shown in Figure 1.9 is used for relative quantification. The peak areas for the heavy (shown in red) and light (shown in blue) isotopes are compared.

In labeled quantification, stable isotopes are used to mark which sample a peptide came from. In this case, the peptides are chemically identical, but the mass shifts caused by the isotopes allows them to be discriminated in a MS. Several different stable isotope labeling techniques have been developed such as: ^{18}O incorporation, Stable Isotope Labeling with Amino acids in Cell culture (SILAC), Isotope-Coded Affinity Tags (ICAT), Isobaric Tags for Relative and Absolute Quantification (iTRAQ) etc. These techniques have different advantages and disadvantages and they can be used for different types of experiments. Stable isotopes labeling can be used for both relative and absolute quantification methods in proteomic studies. I will mostly focus on relative quantification methods since they are related to my work.

1.6.2.1. Label-Free Quantification

Label-free quantification is used to determine the relative amount of proteins in two or more samples. Label-free quantification methods can be mainly divided into two categories; spectral counting and signal intensity measurements.

1.6.2.1.1. Relative Quantitation by Peak Intensity (MS-signal Intensity)

Peak intensity is based on the integrated measurement of chromatographic peak areas for any given peptide in an LC-MS runs. Importantly, the measured peak intensity is proportional to the concentration of the measured peptide [61]. After the sample has been analyzed and peptides identified, the ion chromatograms for every identified peptide are extracted from the LC-MS/MS run and their peak areas are integrated over the chromatographic time scale [62]. These extracted ion chromatograms (XIC) are used to

determine the relative abundance of each peptide. The measurement is basically done by comparing peak areas for the same peptide in different samples. However, peak areas are compared only if they fall within the same *m/z* and retention time windows. Quantification is based on the MS-response of the precursor ion. Therefore it is important that the same peptide is identified in every run, otherwise the quantification becomes inaccurate.

The main problem of this approach is the experimental variations between samples. Since each sample and replicate is processed and run independently, even small technical errors can affect the resulting MS-signal intensities of peptides. Normalization with internal standards and/or normalization coefficients are typically used to correct for these errors and provide accurate quantification.

1.6.2.1.2. Relative Quantitation by Spectral Counting

Spectral counting is based on the idea that the number of peptides sequenced for a given protein is correlated to its quantity. Accordingly, quantification involves comparing the number of identified spectra of the same protein in different samples [63]. This method requires proteins to be identified by a sufficient number of spectra to be accurately quantitated. In this approach, the more abundant proteins produce more MS/MS data than the less abundant ones, so more abundant proteins are sample more frequently by the MS than low abundance peptides. Therefore, this method favors the quantification of the higher abundance proteins, while the quantification of low abundance proteins can be problematic, especially if the identification is based on only one or two spectra.

In contrast to peak intensity methods, which requires specialized algorithms to align and compare LC-MS peaks, no specific tools or algorithms are needed for spectral counting due to its easy implementation. The data is typically normalized to the number of peptides a protein produces [64]. Since large proteins tend to contribute more spectra (since they

have more peptides) than small proteins, a normalized spectral abundance factor (NSAF) is used to account for the effect of large proteins on spectral count. Furthermore, an additional step of normalization to the total number of spectra identified in sample is also frequently used to normalize for technical errors that effect the overall quality of the whole data set [64].

The disadvantages are very similar to label free quantitation based on ion intensity measurements. Since each sample is run separately, sample to sample variability needs to be addressed, and similarly to the ion intensity based studies, internal standards play an important role.

In general, label-free techniques work really well on samples that are mostly the same, because large variations in the expression of many proteins, will affect the quantification of co-eluting peptides. This is principally due to the fact that all the peptides that enter the mass spectrometer at the same time are competing to be ionized and for mass spectrometer acquisition time. For example, the loss of a highly abundant peptide will result in new peptides being selected for MS/MS. Over the course of a run this is usually not an issue so long the samples being compared are similar. Label-free techniques are frequently used because they do not required expensive reagents or specialized tools.

1.6.2.2. Labeled Quantification

Labeled quantification involves the incorporation of stable isotopes into the sample. The isotope labeling takes advantage of the fact that the same peptide, which has incorporated different isotopes will be chemically identical, but will have different masses due to the presence of heavy isotopes in one of the peptides. This allows samples labeled with different stable isotopes to be mixed together and analyzed in the same LC-MS/MS run. The two samples will undergo the exact same separation, the exact same competition for

ionization, have the exact same chance to be acquired for MS/MS etc. The relative amount of each peptide in each sample is detected in the spectra because there will be a characteristic mass shift depending on the isotopes used.

Labeling techniques can be divided into three groups; metabolic, chemical or enzymatic. Metabolic labeling involves labeling *in vivo* during growth by adding isotope labeled amino acids in cell culture medium or in food. On the other hand, the chemical approach usually involves covalently labeling the proteins after they are extracted from the cells. Finally, the enzymatic labeling is performed during enzymatic digestion. Any stable isotope (heavy) can be used as tags, such as carbon-13 (^{13}C), oxygen-18 (^{18}-O), nitrogen-15 (^{15}N) or deuterium (^{2}H). The naturally occurring light isotopes for these elements are: ^{12}C, ^{1}H, ^{14}N, ^{16}O [65]. However, the elements that are not commonly found in proteins typically are not used for metabolic labeling.

1.6.2.2.1. Chemical Labeling

In the chemical labeling approach, the sample is not limited to living systems, and it is applicable to all types of biological samples. Chemical or enzymatic labeling can be performed using a variety of isotopically labeled chemicals which are covalently attached to the proteins or peptides. Typically the amino, carboxyl, or sulfhydryl groups in proteins are targeted for labeling.

Labeling of the amino groups is based on the idea that, proteins can be covalently labeled on the amines found at the N-terminus of peptides and on the side chains of lysine and arginine. These amines are typically targeted because the chemistry involved is typically robust and specific. The *isotope-coded protein label* (ICPL) and the iTRAQ labeling approach both involve the labeling of amino groups and both can be used for multiplex (comparing 4 to 8 samples at a time) analysis by varying the number of stable isotopes in the chemical

label [66]. iTRAQ labeling has the added advantage of being isobaric. In this case the overall mass of the label is the same for all isotope forms, as it contains an isotope labeled group and a balancer group. The isotopically labeled group is only detected when it is split off from the balancer during MS/MS acquisition and this allows for multiplexing without increasing the complexity of the sample [67].

Besides labeling amino groups, labeling of the thiol groups in cysteines is also commonly used, and can also be used to reduce sample complexity. Since it is related to my work, I will describe a specific method where thiol groups are tagged, ICAT.

1.6.2.2.1.1. Isotope Coded Affinity Tags (ICAT)

ICAT is a relative quantitation method that was introduced by Aebersold and co-workers in 1999 [68]. In this technique, the thiol groups of cysteine are labeled with either light or a heavy tags using a custom designed alkylating reagent.

ICAT reagents consist of three general elements: (I) a reactive group capable of labeling sulfhydryl groups (e.g., iodoacetamide to modify cysteine residues), (II) an isotopically coded linker, which can contain multiple, ^{12}C, ^{1}H, ^{13}C, or ^{2}H atoms to give an 8 Da mass difference, (III) and a biotin tag, which is used for the affinity isolation of labeled proteins. Labeling can be performed on proteins or peptides. The mass shift observed in the mass spectra depends on the linker. A linker with 8 hydrogen/deuterium would give a ~8 Da mass difference.

For the quantitative comparison, one sample is labeled with an isotopically light probe and the other with an isotopically heavy probe. After labeling, the samples are combined, digested with a protease and subjected to avidin affinity chromatography to enrich for the labeled cysteine containing peptides **(Figure 1.10)**. These peptides are then analyzed by LC-MS/MS [69]. In the same mass spectra, the peptide pairs are seen as doublets and their relative intensity reflects the relative abundance of the corresponding protein in both samples.

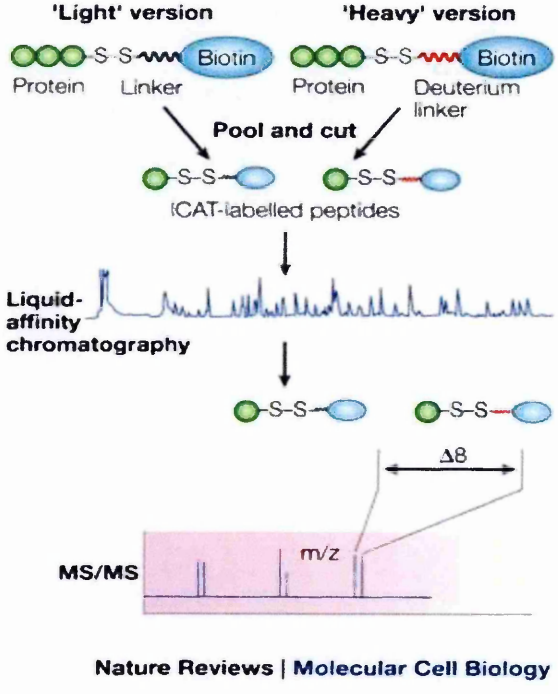


Figure 1.10. Workflow for ICAT experiments [256]. ICAT involves the chemical labeling of cysteine residues with an isotopically coded affinity tag. The cysteine containing peptides are purified using the tag and then analyzed by MS. The peak intensities of the isotope pairs (*Δ8*) are used for relative quantification.

The principle advantage of ICAT is that, it only labels a subset of peptides, in this case peptides with cysteine residues, which results in reduced sample complexity and potentially more protein identifications. However, these advantages are offset by a number of disadvantages. First of all, cysteine is a relatively rare amino acid. Since most

proteins have only 1 or 2 cysteines, quantitative data becomes weak due to lower sequence coverage. Moreover, there are many intervening steps between cell lysis and sample mixing and these steps invariably result in systematic errors being introduced between the samples.

One of the experimental steps involves the use of Biotin: Avidin affinity chromatography. Although this step is performed after sample mixing, it creates a variety of problems. The strength of this interaction results in significant sample losses, which greatly affects sensitivity and naturally occurring biotinylated peptides also interfere with analysis. Several techniques have been improved for removal of biotin tags such as using acid-cleavable and photocleavable biotin tags [70,71]. However, ICAT has been supplanted by other methods, primarily because, even with these improvements, the workflow is too labor intensive.

1.6.2.2.2. Enzymatic Labeling

^{18}O-labeling, which is also known as Enzyme Mediated Oxygen Substitution (EMOS), is the principle form of enzymatic labeling. This labeling is performed during enzymatic digestion by performing the digestion in ^{18}O labeled water. Digestion with trypsin results in the incorporation of one or two ^{18}O at the C-terminus of the peptide. The presence of one or two labels at the end complicates the analysis and even with two labels of ^{18}O, many mass spectrometers do not have the resolution to completely separate the isotopic peaks from multiply charged peptides [72].

1.6.2.2.3. Metabolic Labeling

Metabolic labeling of peptides has a long history. Although, the incorporation rate of radioactive isotopes is typically rather low, the incorporated radioactivity can be detected

with great sensitivity, especially when the sample is separated by 1D or 2D gel electrophoresis [73]. However, coupling isotope labeling to MS detection and quantification, the rate of isotope incorporation becomes a critically important factor.

There have been a limited set of studies where the precursor for metabolic labeling is not an amino acid. For example, using metabolites such as (^{13}C)-glucose or (^{15}N)H_4Cl as the carbon and nitrogen sources. This type of labeling will incorporate a variable number of atoms depending on the length of the peptide and the number of specific atoms that each peptide. As a result, due to the lack of uniformity, data analysis is quite complex. However, using isotopically labeled amino acids as a metabolic precursor is much more straightforward and reproducible.

As stable isotope labeling techniques developed, new names have been used such as, stable isotope labeling with amino acids (SILAA) and stable isotope labeling with amino acids in cell culture (SILAC) [74]. These labeling techniques use the metabolic activity of living cells to incorporate stable isotopes and can be applied to tissue culture cells or even living plants and animals. However, metabolic labeling of tissue culture cells is typically used for quantitative proteomics as labeling whole organisms is typically too expensive. The mass shifts produced by the labeled amino acids can be carefully selected and it is even possible to multiplex as amino acids with differing numbers of isotopes can be selected. In addition to quantification of protein levels, SILAC has been adapted for other purposes, including monitoring post translational modification, such as acetylation.

1.6.2.2.3.1. Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC)

SILAC is a widely used quantitation technique for mass spectrometry studies because it is a relatively simple and straight forward approach for MS based quantitative proteomics

[75]. SILAC utilizes the metabolic incorporation of a 'heavy' or a 'light' version of specific amino acids into proteins.

In the SILAC method, the metabolic labeling is performed during protein synthesis in living cells, and growth medium is prepared where natural ('light') amino acids are replaced by 'heavy' SILAC amino acids **(Figure 1.11.)**. After a number of cell divisions, the label becomes incorporated to a very high extent, as cell division and protein turnover depletes the non-isotopically labeled precursors from the sample. Like other stable isotope labeling techniques, SILAC results in the production of two chemically identical groups, which can be separated by the presence of either heavy or light isotopes. The cells grown in the two different media are combined and further treated as a single sample. Since one sample contains light isotopes and the other contains heavy isotopes, the relative amount of proteins is determined by the ratio of the signal intensities of peptide pairs (light and heavy).

Once the metabolic incorporation is complete, either the heavy or light enriched group can be perturbed and allowed to recover and then the two groups are combined prior to cell lysis. The early mixing of the two samples minimizes the systematic errors that often creep into quantitative experiments. The mixed sample is digested with a protease and then analyzed by mass spectrometry. Mass difference from the stable isotope labeling amino acids are measured. Typically peak areas from the precursor mass of heavy and light peptide are used to perform the quantification [76]. Changes in protein level between the two groups are indicative of the cellular response to the perturbation.

In principle, any amino acid can be used for metabolic labeling. However, all the amino acids should be abundant in the medium to avoid de novo amino acid synthesis in cells, which can scramble the isotope labeling [77]. Leucine, lysine, tyrosine and arginine are the most used amino acids for SILAC [77]. A double labeling, using Arg/Lys ($^{12}C_6/^{13}C_6$-arginine/$^{12}C_6/^{13}C_6$-lysine) is frequently used, because all of the peptides in a sample can be labeled if trypsin is used for the digestion.

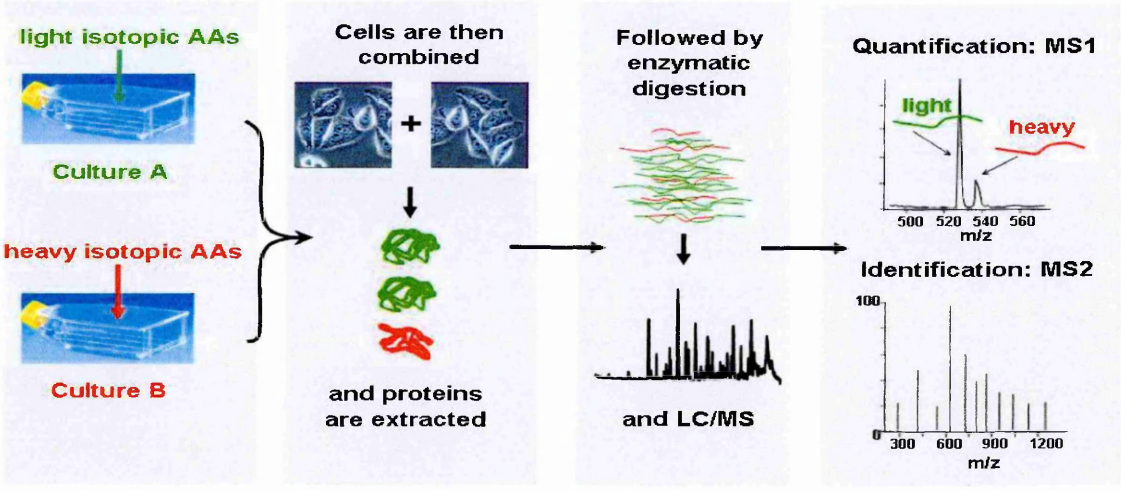


Figure 1.11. Workflow for SILAC experiments [78]. Cultures are grown in the presence or absence of isotopically labeled amino acids and then combined. Following sample preparation and analysis by mass spectrometry, the peak intensities for the isotope pairs are used for relative quantification.

1.6.2.2.3.1.1. Adaptations of SILAC

Comparative proteomics is the most common application of stable isotope metabolic labelling. However SILAC can be used for different studies such as, quantification of PTMs and measurement of protein turnover rates.

The study and quantification of PTMs can be performed with isotope labeling using two different strategies: by labeling the peptide or by labeling the PTM. The first approach involves stable isotope labeling all of the peptides and purification of the peptides carrying

PTM of interest. The PTM itself is not labeled in this category. The second approach involves direct targeting of PTM and their replacement by reporter group [79]. However, direct labeling of PTM is distinct from strategies for labeling proteins or peptides.

The use of SILAC for quantification of several protein modifications was demonstrated in several studies. One of them is heavy methyl SILAC. This method allows identification and simultaneous quantification of protein methylation sites, by labeling the peptide. In this method, since quantification in SILAC is based on individual peptides, site-specific quantification of protein methylation is achieved by determining the intensity ratios of methylated peptide pairs. These ratios can be normalized against the protein ratio obtained from unmethylated SILAC-labeled peptides [80].

The SILAC approach also used for more comprehensive analysis of several protein modifications. One group studied the dynamics of histone modifications such as, acetylation, methylation and phosphorylation, throughout the cell cycle [81] . Another group used SILAC coupled to a histone peptide pull-down approach to screen specific interactors of histone H3 trimethylated on Lys-4 (H3K4me3). Same group used triple SILAC pull-down assays to further analyze these interactions [82].

The SILAC approach was also used for Sirt3-regulated acetylome analysis [83]. Acetylated peptides were enriched from the sample by pull-down assays and SILAC ratios were determined. The SILAC technique was also used in several other studies for mapping the acetylation sites [84,85].

Ubiquitination studies are another area where SILAC approaches have been used. Large scale quantification of ubiquitination sites using SILAC has the advantage of simultaneously identifying and quantifying site specific ubiquitination [86]. Quantitative analysis of

ubiquitination has also been studied in HeLa cells by using tandem affinity purification of ubiquitinated proteins from cell lysates [87].

Another PTM study that can be done with SILAC approach is the analysis of phosphorylation. Phosphopeptides are enriched using immobilized metals or metal oxides (IMAC, MAOC) or using antibodies and the sites of phosphorylation were quantified by amino acid labeling using SILAC [88]. In this case, there is only one stable isotope for phosphorous, so it is impossible to directly target the PTM.

SILAC can also be used for *in vivo* animal experiments [89,90] and it is known as stable isotope labeling of mammals (SILAM). This is performed with a diet that is composed of stable isotope-enriched proteins. Although, this ensures that all the cells from these animals will contain labeled proteins, the procedure is lengthy, since it requires at least two generations of labeling, and is quite expensive. The TTMSILAC technique was introduced to help overcome the expense of SILAM [91]. This method involves adding a mixture of lysates from several different SILAC labeled cell lines as an internal standard to the samples to be quantitated. It has mostly been used for comparison of cell lines to patient tumor samples.

In the 'culture-derived isotope tags' technique which derived from SILAC [92], a labeled internal standard is used for quantifying the proteome of a given tissue. This enables using SILAC to quantify changes in samples that are too difficult to label, such as in human tissue samples.

These studies show that metabolic labeling approaches, especially those based on SILAC are a powerful tools to qualitatively and quantitatively study complex proteomes.

1.6.2.2.3.1.2. Advantages and Limits of SILAC

SILAC has several important advantages. SILAC labeling is very efficient and the downstream sample processing and mass spectrometry protocols are virtually identical to standard, non-quantitative workflows. Up to five different cellular states can be compared using a multiplex SILAC-based strategy [93]. However, due to the nature of the isotope labeling, SILAC greatly increases the complexity of the sample entering the MS. The increased sample complexity may result in fewer overall protein identifications, especially when multiplexed.

Although, SILAC increases the sample complexity relative to label-free methods, the quantification is particularly accurate. This is largely due to the early mixing of the samples and the relatively straightforward sample processing steps.

1.7. The Novel Hybrid Method

In this thesis, I aimed to develop and evaluate a method for mass spectrometry based proteomics. The main goals were to improve the rate of protein identification and quantification.

The first step in my project was to develop a hybrid quantification protocol that has the advantages of both SILAC and ICAT. The hybrid approach I developed uses classical SILAC labeling coupled to a "novel" scheme for enriching cysteine containing peptides. This enrichment step resembles ICAT since cysteine and cystines are targeted for modification and enrichment. In the hybrid protocol I developed **(Figure 1.12.)**, this modification involves oxidizing the cysteine/cystine residues to cysteic acid and selective enrichment of these peptides based on the properties of cysteic acid. In brief, I aimed to

integrate the advantages of these two techniques into one simple, straightforward technique.

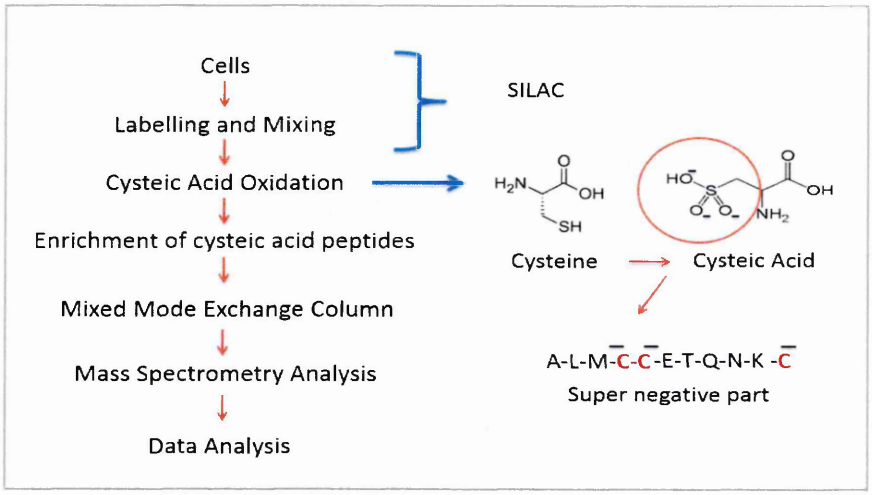


Figure 1.12. Workflow of Hybrid Method. The various steps of the hybrid workflow are shown.

For cysteic acid enrichment, I adapted a chromatographic technique called, 'Electrostatic Repulsion and hydrophilic Interaction Chromatography' (ERLIC). ERLIC affords convenient separations of highly charged peptides that cannot readily be resolved by other means. There are two different modes for performing ERLIC. It can be used with cation exchange columns (PolySULFOETHYL A or PolyCAT A) for positively charged analytes or it can be used with anion exchange columns (PolyWAX LP) for negatively charged analytes, such as cysteic acid containing peptides. ERLIC chromatography is based on superimposing electrostatic effects on hydrophilic interactions and this selects or de-selects for specific functional groups. Therefore, the influence of ionic groups can either be minimized or maximized, thereby altering the degree of retention. Ideally, in the hybrid method only peptides with cysteic acid will strongly interact with the positive surface of the resin, resulting in a quick and efficient enrichment of cysteic acid containing peptides **(Figure 1.13.)**.

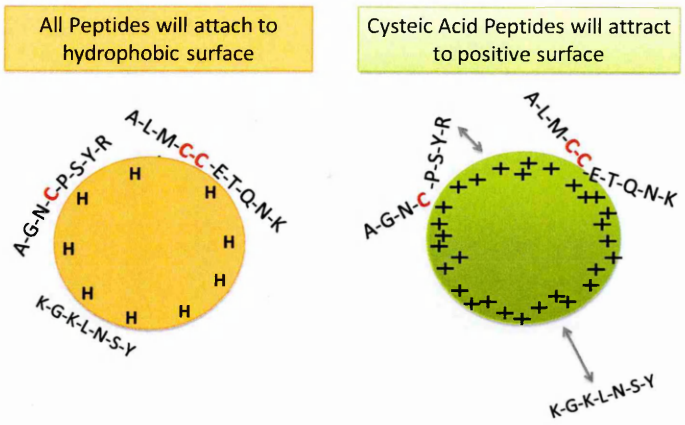


Figure 1.13. The selection of cysteine containing peptides with PolyWAX LP column by ERLIC chromatography. ERLIC chromatography can be used to select for certain peptides, and in this case positive charges are being repelled.

1.7.1. Advantages and Limits of Hybrid Method

The Hybrid technique uses the classical SILAC technique, coupled to a novel scheme for enriching cysteine containing peptides. Hence, the hybrid technique possesses all the advantages of the SILAC technique itself.

On the other hand, since ICAT only labels approximately 2-5% of the peptides, it is reducing the sample complexity. However, instead of the complex chemistry of the ICAT reagents, I am using a much more streamlined chemistry, where performic acid is used to convert cysteine and cystine into cysteic acid. This results in the creation of a "super negative" side chain that can be used for selective enrichment of these peptides **(Figure 1.12.)**. As a result, I get the ease of labeling and enhanced quantitative performance of SILAC protocols with that added benefit of the reduction in sample complexity by cysteine enrichment as seen with ICAT protocols. Moreover, while ICAT uses only cysteine-containing peptides for quantification, the unbound peptides from the ERLIC chromatography can be collected, analyzed and quantitated. Since, the hybrid method

uses all the fractions (cysteine-containing peptides, and the others), it quantitates a much larger set of peptides and proteins.

The initial optimization and eventually was performed using BSA. BSA contains 35 cysteine residues out of some 580 total residues, or roughly means nearly 6% of BSA is cysteine residues, which makes it ideal protein for these optimizations. After the development of method is finish, I moved the second part of the thesis.

In the second part of my study, I used the hybrid protocol I developed, to study a biological system. I decided to use my protocol to study the effects of a cancer therapeutic compound, sirtinol, on a tumor cell line. Chemically induced models are easy to reproduce and reproducibility is an important factor for proteomic studies, especially when evaluating new protocols. Importantly, sirtinol is able to induce senescence in many tumor cell lines, and this system should also provide insights into cellular senescence and therapeutic compounds that activate this response.

1.8. Pharmacological Induction of Senescence as a Cancer Therapy

Cancer is one of the most important causes of death around the world. A number of human malignancies can be cured with current therapies and prolonged survival can be obtained. However, there are many situations where current therapeutic approaches are not sufficient, thus there is still a great need for preclinical studies and further drug development.

A promising approach for cancer treatment is permanently disabling the proliferative capacity of cancer cells without inducing the cell death. This can be achieved by therapy-induced senescence. It has been more than a decade since it was initially suggested that cellular senescence may be one of the mechanisms by which cancer chemotherapy drugs

work in vivo [98]. Many chemotherapeutic agents cause DNA damage and induce senescence in normal cells. Interestingly, it has been shown that they can also induce senescence in tumor cells. Anticancer chemotherapeutic agents and ionizing radiation have also been shown to induce senescence-like growth arrest in human cancer cell *in vitro* and *in vivo* [94–97]. The induction of senescence was verified by the induction of SA-β-gal staining *in vivo* in rodents [95,98,99]. Moreover, it has been observed in patients with breast cancer [96].

Genetic abnormalities are one of the major hallmarks of cancer. However, studies over the past decades indicate that epigenetic regulation is also involved in the development of cancer. DNA methylation and histone modifications are the major epigenetic factors that are frequently altered in tumor cells. Acetylation of histones is controlled by histone acetyltransferases (HATs) and histone deacetylases (HDACs). HDACs are a group of enzymes that catalyze the removal of acetyl groups from both histone and non-histone proteins. They are involved in modulating many key cellular processes such as transcriptional regulation, DNA damage repair, apoptosis, cell cycle progression, metabolism and senescence.

The functional activity of the HDACs varies depending on their structure and intercellular localizations. There are four classes of HDAC. Class III HDACs are more generally called sirtuins, which were named after the yeast gene 'silent mating-type information regulation 2'. In mammalian systems, sirtuins are divided into five subclasses which have different enzymatic activities. Sirtuins have been shown to be the master regulators of cellular senescence[100].

Currently, histone deacetylase (HDAC) inhibitors are under intensive investigation as cancer therapeutics, since they have shown promising effects for various cancers. A major barrier for HDAC inhibitors is that they have cytotoxic effects on normal host cells [101].

In this thesis, I used a SIRT1-specific HDAC inhibitor, sirtinol, to treat H1299 cancer cells. Importantly, sirtinol induces a senescent-like growth arrest in these cells. Since senescence is a potential tumor suppressor mechanism, sirtinol is a promising drug candidate. However, the aim of this study was mostly elucidating the proteome changes in response to sirtinol, rather than characterizing its efficacy as a cancer therapeutic drug. My work will give a better understanding of the mode of action of sirtinol, as well as revealing the long-term biological responses to sirtinol treatment.

The causes of senescence and its relationship with cancer will be explained broadly in the next sections.

1.8.1. What is Cellular Senescence?

Cellular senescence was originally defined as 'irreversible cell cycle arrest' caused by 'replicative exhaustion' in human diploid fibroblasts (HDFs) [102]. Experiments showed that fibroblast in culture divide robustly when first placed in culture. However, after many cell doublings, their proliferation rate declined [102]. This suggested that all cells in culture lose the ability to divide, even when provided with enough space and nutrients. This led to two important hypothesis about senescence. The first hypothesis was that senescence could be an anti-cancer or a tumor-suppressive mechanism, since many cancer cells proliferate indefinitely in cell culture [103]. In this respect, senescence is considered beneficial for an organism, since it puts an end to proliferation of cancer cells. The second hypothesis, on the other hand, considered senescence a deleterious response, based on the fact that tissue regeneration and repair deteriorates with age and is correlated with

an increase in senescence. However, it is clear that many types of stimuli can induce senescence such as, DNA damage, oncogene activation, and physiological stress.

1.8.2. Characteristics of Senescent Cells

In multicellular organisms, such as mammals, there is typically a mix of both mitotic and post-mitotic cells. There are some characteristics of senescent cells that can be used to discriminate them from normal post-mitotic cells. Irreversible growth arrest, resistance to apoptosis and altered gene expression are the major hallmarks of the cellular senescence.

Growth arrest is the main hallmark of the cellular senescence. Although mitotic cells can proliferate, they can also be in a state which called 'quiescence,' or G0, where they reversibly stop dividing. Quiescent cells can start dividing in response to appropriate signals. On the other hand, post-mitotic cells (i.e. neurons, myoblast and osteocytes) and senescent cells permanently stop dividing. Senescent cells usually have a DNA content that is typical of G1 phase, but they remain metabolically active. Once a mitotic cell enters senescent growth arrest, it cannot start DNA replication, even in adequate growth conditions. This replication failure is primarily caused by the expression of cell-cycle inhibitors.

Together with growth arrest, senescent cells undergo a morphological change and are generally enlarged, often doubled in volume and they adopt a flattened morphology. Histochemical staining for senescence-associated Beta-galactosidase (SA-βgal) is a commonly used marker for senescent cells [104]. The SA- βgal activity is produced by the lysosomal β-galactosidase, which reflects the elevated lysosomal biogenesis that frequently occurs in senescent cells [105].

Another characteristic of senescent cells is the resistance to apoptosis. Apoptosis is a process of programmed cell death, which occurs in many organisms. Like senescence, apoptosis is an extreme response to cellular stress and it is considered one of the most important tumor-suppressive mechanisms. However, while senescence prevents the growth of stressed or damaged cells, the cells remain metabolically active, while apoptosis completely eliminates the cells [106]. Interestingly many senescent cells are resistant to apoptotic signals. For example, senescent human fibroblast resist apoptosis caused by oxidative stress and growth factor deprivation [107], however they do not resist apoptosis caused by the Fas death receptor [108]. This resistance to cell death of helps explain why the number of senescent cells increases with age and why they are so stable in culture.

Another feature of senescent cells is altered gene expression, including the genes for cell-cycle inhibitors or activators [109]. In senescence cells, the levels of cell cycle inhibitors increases and the genes promoting cell cycle progression decrease. For example, the expression of the tumor suppressor protein p16 is low in most normal cells, it is readily detectable in senescent cells [110]. Moreover, it is interesting that many gene expression changes do not seem related to growth arrest but they are mostly related to changes in the micro-environmental state of the tissue [111]. For example, senescent fibroblast overexpress proteins that remodel the extracellular matrix. Given the fact senescent cells increase with age, it is possible that they might be responsible for the age dependent alterations in tissue structure and function.

Some senescent cells can also be identified by the presence of senescence-associated DNA-damage foci (SDFs) [112-114]. SDFs are present in senescent cells (both human and mouse cells) and contain proteins that are associated with DNA damage, such as phosphorylated histone H2AX and p53-binding protein-1.

Senescence-associated heterochromatin foci (SAHFs) are also considered to be a cytological marker for senescence. SAHFs preferentially bind to DNA dyes such as DAPI. However, since pericentromeric foci in mouse cells are much more prominent than in human cells, they can be mistaken for SAHFs.

1.8.3. Senescence-Associated Secretory Phenotype (SASP)

SASP is the final feature of senescent cells. This feature is considered important since it explains the role of cellular senescence in organismal aging [111,115]. SASP components include growth factors, proteases, cytokines and chemokines [111,116]. There are a myriad of biological activities associated with this phenotype. Some SASP factors are known to be beneficial for senescent cells, while others have deleterious effects.

SASP, together with its biological activities, can be considered a very complex system. In some situations, depending on the physiological context, SASP can stimulate or inhibit WNT signaling and cell proliferation. Chronic WNT signaling can cause senescence in both differentiated cells and in stem cells [117].

Related to senescence and age-associated diseases, SASP factors such as IL-6, IL-7, MCPs (monocyte chemoattractant proteins), and MIPs (macrophage inflammatory proteins) directly and indirectly promote inflammation. Secretion of these proteins is assumed to cause chronic inflammation locally and perhaps systematically [116]. It's known that chronic inflammation is involved in many age-related, degenerative and hyperplastic diseases [118].

1.8.3.1. What Causes SASP

Normal cells that senesce due to overexpression of p21 or $p16^{INK4A}$ do not express SASP, but they go under growth arrest and show other characteristics of senescent cells. However, if the cells senesce due to DNA damage, dysfunctional telomeres, epigenetic

disruption, oxidative stress, mitogenic signals and other senescence inducers, they develop SASP of varying types [119]. The function of SASP may be for communication with neighboring cells to prepare the tissue for repair or to aid in the clearance of damaged cells by the immune system.

Many of the SASP components are positively regulated by DDR proteins such as ATM and CHK2 (checkpoint kinase 2) [120,121]. These proteins activate p53, which normally acts to restrain the SASP. It should be noted that DDR induces SASP, only when it is persistent. How DDR signaling promotes the expression of SASP components is not very well known.

1.8.4. Causes of Cellular Senescence

There can be many triggers for senescence such as dysfunctional telomeres, DNA-damage, chromatin perturbation and stress. Regardless of the type of trigger, the establishment of the senescent phenotype requires different effectors, but once established the phenotype is remarkably similar. For example, replicative senescence corresponds to a response to modest but long-term exposure of cells to the stress that occurs from many rounds of proliferation, while oncogene-induced senescence is a more acute and dynamic process. How these different conditions trigger similar pathways involved in senescence is a matter of great interest **(Figure 1.14.)**.

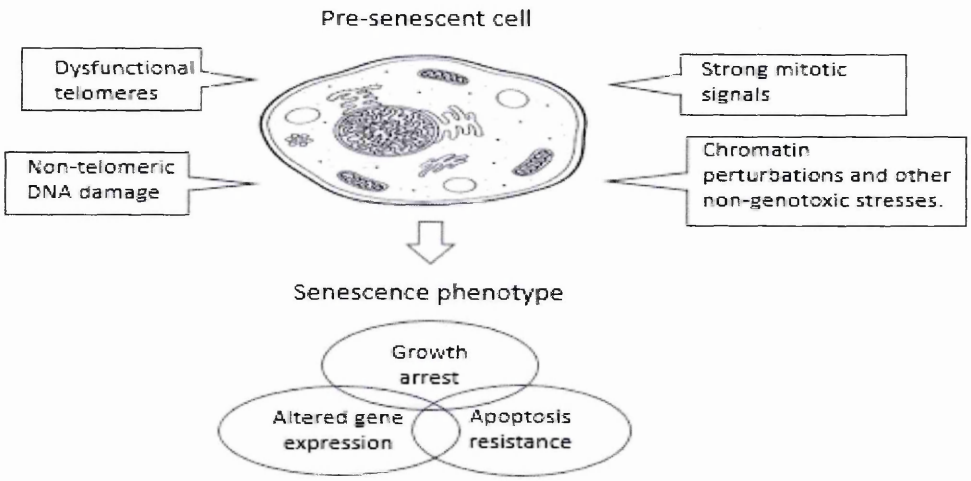


Figure 1.14. Different factors that can trigger senescence.

1.8.4.1. Telomere-dependent Senescence

Telomeres have a major role in cellular senescence and the shortening of telomere puts a limit to cellular replicative capacity. Telomere shortening was one of the first causes for the cessation of cell proliferation [122,123]. Normal cells typically lack the enzyme, telomerase, which functions to lengthen the telomere after each round of cell division. In the absence of telomerase, the telomere get successively shorter. When telomeres become sufficiently short, cells enter in a growth arrest called senescence. However this situation only applies normal cells. Telomerase is often found in cells with a larger replicative potential, such as stem cells, reproductive cells, and cancer cells [124]. This places a limitation on the cellular renewal of most tissues and telomere shortening is believed to a major contributor to the aging phenotype.

Telomeres consist of repetitive nucleotide sequences (5'TTAGGG-3' in vertebrates) and these repeats are associated with proteins that form a chromatid cap that stabilize the end of the chromosomes. The chromatid cap protect chromosomes from degradation

and fusion with neighboring chromosomes [125]. In each round of cell division, the linear DNA becomes shorter than its template, since DNA polymerases cannot replicate DNA ends completely. This phenomenon, 'the end-replication problem', causes 50-200 base pairs of telomeric DNA loss per round of replication. Since telomeres range from 10-15 kb in humans, only a certain number of cell divisions are possible before the telomeres become critically short and dysfunctional.

Dysfunctional telomeres trigger a classic DNA-damage response (DDR) which enables cells to sense the damaged DNA, specifically as double-strand breaks (DSBs) [112,113]. Response to DSBs can be either to repair the damage, apoptosis or entry into cell-cycle arrest. Importantly, senescent cells respond to telomere shortening with a prolonged DDR activation. Although the severity of the damage is an important factor, how cells choose between transient DDR activation and the persistent DDR signaling seen in senescent cells is not quite clear. Moreover, it has been shown that even a few short telomeres are enough to trigger senescence [126,127].

Several proteins are involved in DDR, such as, protein kinases (i.e., ATM, DNA-PKcs), adaptor proteins, and chromatin modifiers such as histone acetylases, deacetylases, histone methylases and demethylases. The protein kinases, ATM and ATR are the central components of the checkpoint mechanisms that signal the presence of damaged DNA or collapsed replication forks. While the ATM/ATR system is the main double strand break signaling pathway, another kinase DNA-dependent protein kinase (DNA-PKcs) also participates DSB signaling. These kinases work together with their regulatory subunits, DNA repair proteins and adaptor proteins to sense the abnormal DNA structures and implement the appropriate DNA damage response. Most of these proteins can also be found in the DNA-damage foci seen in senescent cells. Moreover, chromatin modifiers

are involved in DDR. Specific histone modifications accumulate at the sites of damaged DNA and facilitate the recruitment of a subset of damage response proteins. For example, phosphorylation of the histone variant H2A contributes the repair by recruiting the sister chromatid cohesion factor cohesin, which is important for efficient DSB repair [128]. DNA damage foci also contain a subset of telomeres in dysfunctional-telomere-induced senescent cells, suggesting that dysfunctional telomeres resemble DSBs.

The length of telomeres can vary in different species. For instance, laboratory mice have very long telomeres (40-60 kb) comparing to human cells (5-15 kb) and moreover they express telomerase, therefore they are more resistant to senesce induced by telomere shortening. Human fibroblasts typically senesce due to rapidly shortening telomeres, while mouse fibroblast in culture typically senesce due to environmental factors. This is often referred as oxidative stress induced senescence.

Telomere shortening does not occur in cells that express telomerase [129]. The numbers and types of telomerase-expressing cells vary widely between species [130]. For example, in mice many cells in the adult animals express telomerase, while in humans, such cells are rare. They are mainly, cancer cells, stem cells and some somatic cells [131,132].

The end replication problem can be prevented by telomerase. This enzyme adds telomeric DNA repeats directly to chromosome ends by using its catalytic component, telomerase reverse transcriptase (TERT) and a template RNA component. Cells normally do not express TERT or they express it at very low levels just to prevent telomere shortening [133]. On the other hand, germ line cells and many cancer cells express TERT. It has been shown that ectopic TERT expression in normal human cells can prevent telomere shortening and the senescence caused by it. However, telomerase cannot prevent senescence that is triggered by other inducers [134].

1.8.4.2. DNA-damage-initiated Senescence

Severe DNA-damage, especially damage that creates DSBs, causes many cell types to undergo senescence [135,136]. Both DNA-damage and telomere initiated senescence depend on p53 and usually involves p21 expression [113,114]. However, in many cells, p16 can also induce cell cycle arrest caused by DNA damage or telomere dysfunction [137,138].

DSBs are very strong inducers of senescence. It has been shown that even a single unresolved DSB can induce senescence [136]. The effective induction of senescence requires persistent DDR signaling after genomic damage. This response is different than a DNA damage response that causes transient growth arrest. Persistent DDR signaling can be identified by the long-term presence of nuclear DNA damage foci.

DNA-damage induced senescence can be seen as a general trigger of the senescence response, especially considering dysfunctional telomeres trigger a DDR. However, since the first trigger is the dysfunctional telomeres, it is considered as a different pathway. Interestingly, dysfunctional telomeres, can be considered among the senescent inducers, which create a DDR without physical damage to the DNA. Chromatin perturbations can also induce a DDR when there is no physical DNA damage. Moreover, by inducing proteins involved in DDR signaling, such as ATR, senescence can be initiated without DNA damage [139].

Many chemotherapeutic agents function by physically damaging the DNA. These drugs can also induce senescence in normal cells. Sometimes this type of senescence is referred to as therapy-induced senescence [140].

Oxidative stress causes DNA damage and single-strand breaks. These lesions need to be repaired or they can be converted into DSBs during DNA replication. Moreover,

oxidative stress can accelerate telomere shortening, presumably because the G-rich telomeric DNA is specifically vulnerable to oxidative damage [141]. When oxidative stress is the initial trigger, this type of senescence is also referred as oxidative-induced senescence.

1.8.4.3. Chromatin Perturbation-Induced Senescence

Chromatin affects the global control of transcription by helping to determine which genes are active (euchromatin) or silent (heterochromatin). This control depends on histone modifications, for example, methylation, and acetylation. It has been shown that, global chromatin relaxation with broad-acting histone deacetylase inhibitors (HDAi) induces senescence by derepressing the $p16^{INK4a}$ gene, which causes the formation of senescence associated heterochromatin [142]. Notably, the $p16^{INK4a}$ tumor suppressor is expressed by many senescent cells and is considered a biomarker for aging [143]. The exact mechanism of chromatin perturbation-induced senescence is poorly understood but it is assumed that it depends on the cell type. For example, in human fibroblasts, HDAi sequentially induces p21 and p16 expression, and the senescent growth arrest depends on the presence of pRb. On the other hand, in mouse fibroblasts, the p53 pathway is more important for the senescent response to HDAi [144]. It is possible that HDAi may cause senescence by initiating a p53-dependent DDR, since it is known to induce ATM kinase activity [139].

It is interesting that HDAi, which can cause reversible disruption of heterochromatin and centromere function[145] also cause senescence. Especially given the role of heterochromatin and senescence-associated heterochromatin foci (SAHF) in establishing and maintaining the senescence growth arrest [142]. Moreover, the down regulation of histone acetyl transferases which promotes heterochromatin formations also induces senescence [146]. It is not known how senescence can be induced by both heterochromatin

formation and heterochromatin disruption. Each manipulation may change the expression of different genes and the response may be cell type specific. Understanding this conflict is important since HDAi are good candidates for certain cancer treatments [147].

1.8.4.5. Stress and Other Inducers of Senescence

Stress-induced senescence or premature senescence (SIPS) is a very broad concept, which includes a variety of stresses and sources of cellular damage. During SIPS the cell is held in growth arrest by combination of gene expression changes, heterochromatin formation and the activation of CDK inhibitors. Although the mediators are different, common pathways are involved in the mechanism of this type of stress.

Importantly, even in a culture of fibroblasts, a single stimuli can induce senescence through different pathways. Some cells in a culture may senesce because of the expression of p16, while others senesce due to p53-dependent DDR [114,148]. The induction of senescence by the p16 pathway is one of the least understood [149]. It has been shown that expression of p16 increases with age in mouse tissues [150] and stem cells *in vivo* [151]. Expression of p16 is controlled by many different factors including HDACs and chromatin formation [144,152]. Moreover, it has been shown that oxidative stress also induces expression of p16 in MEFs [135].

Persistent DDR signaling initiates senescence growth arrest. This signaling is also accompanied by the activation of other signaling pathways including, p38MAPK and the protein kinase C pathways [153]. Although it is not known how these pathways are initiated, these additional pathways also stimulate the expression of p16, which acts through pRb and ensures the irreversible growth arrest [154].

Oncogene Induced Senescence (OIS) is another major type of senescence that is induced by oncogenes and other drivers of cell proliferation. Senescence studies with tissue culture models of OIS have proven very useful for characterizing senescence effector mechanisms, and provide a framework for understanding the cancer-senescence connection [155,156].

1.8.5. Cancer-Senescence Connection

Starting from telomere dysfunction, DNA damage, and chromatin disruption, senescence can be triggered by many stimuli. All these stimuli have one thing in common. All of them have the potential to promote cancer. As stated earlier, normal human and mouse cells can senesce in response to intense mitogenic signals. Thus it follows that senescence can be a mechanism for tumor suppression by irreversibly arresting the growth of cells which are at risk for turning into tumors.

Cancer cells have deficits in the regulatory circuits that govern normal cell proliferation and homeostasis. Most cancer cells have six essential alterations in cell physiology that collectively lead to malignant growth. First of all, cancer cells must obtain self-sufficiency in growth signals. Similarly, they must become insensitive to growth-inhibitory signals and obtain limitless replicative potential. Cancer cells also have to evade apoptosis, and be able to sustain angiogenesis. Each of these physiological changes is necessary for tumor development. And each step also represents an anti-cancer mechanism that must be overcome by the mutation of key regulatory proteins.

Acquisition of the six main capabilities during the course of tumorigenesis is linked directly to changes in the genomes of cancer cells. However, there are a number of efficient systems to reduce the mutation rate and maintaining genomic stability. These systems work to ensure the DNA sequence remains intact. These systems ensure that mutations

are rare and that multiple mutations in key genes are highly unlikely to occur at the same time. Still, cancer cells acquired increased susceptibility to mutations [157] that leads to the tumorigenic phenotype. This paradox is explained by malfunctions in the systems that are responsible for genomic stability. The most prominent member of these systems is the p53 tumor suppressor protein, which induces a cell cycle arrest to allow for DNA repair to take place and can induce apoptosis if the damage is irreparable. It is now known that the functioning of p53 DNA damage signaling pathway is lost in most of the human cancers [158]. Moreover, an increasing number of other genes which are involved in sensing and repairing DNA damage or providing correct chromosomal segregation during mitosis have been found to be lost in different cancers. This is one reason why they are referred to as tumor suppressors [159]. Their loss of function is thought to produce genomic instability and eventually the generation of mutant cells with acquired advantages.

Most normal cells needs growth factors to pass from quiescence to active proliferation. Tumors arise from the abnormal proliferation of a subset of cells which do not respond to normal signals that limit cell division. Hence, cells have clever defense mechanisms to tumors, primarily based on blocking the expansion of these abnormal cells. Two major mechanisms are involved in this defense; apoptosis and senescence. However, apoptosis may fail if the complete elimination of tumor cells is not achieved. On the other hand, cells can follow a different defense mechanism; senescence.

It has been found that certain oncogenes, genes that have the potential to cause cancer, or the loss of anti-mitogenic tumor-suppressor genes can also induce senescence in normal cells [160]. This type of senescence is known as oncogene induced senescence (OIS). This mechanism was first discovered when an oncogenic form of RAS was expressed in normal human fibroblasts [161]. RAS chronically stimulates the Mitogen-Activated Protein

Kinase (MAPK) signaling pathway and induces senescence in normal cells [161]. Although the initial response to RAS is abnormal proliferation, the cells sense this abnormality and undergo senescence as a delayed response to counter these oncogenic signals [162]. Later on, other members of the RAS signaling pathway including RAF, MEK, MOS and BRAF as well as, pro-proliferative nuclear proteins such as E2F1 were shown to induce senescence when overexpressed [163–165]. The senescence of cultured human fibroblasts can be avoided by disabling their pRb and p53 tumor suppressor proteins, enabling these cells to continue multiplying for additional generations until they enter into a second state termed crisis. This crisis state is characterized by massive cell death, karyotypic disarray associated with end-to-end fusion of chromosomes, and occasional emergence of a variant (1 in 10^7) cells that has acquired the ability to multiply without a limit, this trait is referred to as immortalization [166].

Most tumor cells that propagated in culture are immortalized, suggesting that the limitless potential of proliferation is a phenotype that tumor cells must acquire. Normal cells have a telomere shortening problem, while cancer cells are able to maintain their telomeres. Most cancer cells do this by upregulating the expression of the telomerase enzyme, which adds hexanucleotide repeats onto the end of telomeric DNA [167]. Even in the absence of telomerase, some cancer cells have evolved another mechanism called ALT, which maintain telomeres through a recombination-based interchromosomal exchanges of sequence information [168].

Since oncogenes induce senescence by stimulating cell growth, excessive mitogenic stimulation can push cells into oncogenic transformation. In fact, it has been shown that when mouse cells are cultured in serum-free medium (normal medium carries high mitogenic pressure), they resist RAS-induced senescence [169].

1.8.5.1. Tumor Suppressors Control Cellular Senescence

In parallel with its role in suppressing cancer, cellular senescence is controlled by several tumor suppressor genes [170]. The most important of these encode the p53 and pRb proteins. Potential cancer cells must lose p53 and/or pRb in order to overcome the proliferative barrier imposed by cellular senescence. This can occur by mutation or epigenetic silencing of one or more important components of the pathways. The pathways controlled by p53 and pRb are crucial for cells to establish and maintain the senescence growth arrest in response to different triggers.

p53 is a transcriptional activator and repressor which controls the expression of genes that cause cell cycle arrest or apoptosis in response to genomic damage. p53 mutations are found in approximately half of human cancers. Not surprisingly, its activity and protein levels increase when cells become senescent [171]. The activation of p53 involves complex set of PTMs such as phosphorylation and acetylation. Since p53 is an important factor in senescence, significant emphasis has been placed on p53 targets, specifically p21, which is elevated in senescent cells and mediates cell-cycle arrest [172].

pRb regulates transcription indirectly by interacting with transcription factors and recruiting chromatin remodeling proteins to genes that control cell cycle progression and differentiation. At the molecular level, if not all many antiproliferative signals are funneled through the pRb and its relatives p130 and p107. When pRb is hypo-phosphorylated, it blocks proliferation by sequestering and altering the function of E2F transcription factors that control the transcription of many genes essential for progression through the cell cycle [173]. Disruption of pRb liberates E2Fs and allows cell proliferation to proceed.

In addition, loss of pRb makes cells insensitive to antigrowth factors that normally operate along this pathway to block the advance through the G1 phase of the cell cycle. For

example, TGF beta, prevents phosphorylation of pRb, which inactivates pRb. In some cell types, TGF beta suppresses the expression of the c-myc gene, which is required for progression through G1 [174]. However, TGF beta can also cause expression of the p15INK4B and p21 proteins, which block the cyclin: CDK complexes responsible for pRb phosphorylation [175].

pRb is largely hypo-phosphorylated in senescent cells. pRb is hypo-phosphorylated, because senescent cells express high levels of p21 and p16 [170]. These proteins inhibit the cyclin-dependent protein kinases, which phosphorylate and inactivate pRb during cell-cycle progression. Elevation of p21 levels is partly because the gene is a direct target of p53 transactivation. While p16 levels increases in senescent cells partly because Ets1, a transcriptional factor that stimulates p16 expression, accumulates in senescent cells. Whereas Id1 levels, which negatively regulates Ets activity, declines [170].

It is clear that cellular senescence entails the activation of several tumor suppressor pathways and inactivation of several oncoproteins that engage both the p53 and pRb pathway. Moreover, ectopic overexpression of p21, p16 also cause growth arrest with a senescent-like phenotype [170].

There are several differences in establishing senescence in human and mouse systems. In mouse cells, inactivation of Rb or p53 is enough to bypass oncogene-induced senescence, while in human cells both pathways need to be inactivated to inhibit OIS [161]. This could be due the longer telomeres found in mice or to the shorter life span of mice than humans, which requires a less robust senescence response to avoid tumors.

Although oncogene-induced senescence does not involve telomere dysfunction, it induces p16 and SAHFs formation like in telomere-induced senescence and many oncogenes induce a robust DDR owing to promiscuous DNA replication. This DDR response is

thought to be the main reason for the initiation and maintenance of the oncogene-induced senescence [160,176].

Senescent cells repress the genes that encode the proteins that stimulate cell-cycle progression, such as, c-FOS, cyclin A, cyclin B and PCNA. Some of these genes are repressed because E2F (transcription factor) induces them, and E2F is inactivated by pRb [142].

In mice, strong oncogenic signals, or the loss of the tumor suppressor protein PTEN, results in the formation of benign lesions which include senescent cells in addition to normal cells [177]. Moreover, benign lesions in human skin which contain cells that express oncogenic BRAF are also senescent [178]. These findings suggest that oncogene-induced senescence can suppress tumorigenesis *in vivo* [179,180].

Loss of these tumor suppressors is essential for oncogenic transformation of human cells in vitro [181], and these pathways are frequently disrupted in human cancer cells [143]. It has even been hypothesized that of aging is a byproduct of the crucial role that senescence plays in blocking cancer.

1.8.6. Potential Benefits of Cellular Senescence

One of the main benefits of cellular senescence is that it suppresses tumorigenesis in vivo. A recurring question in cellular senescence is why cells that are susceptible to cancer do not simply choose apoptosis, which does not involve complex responses such as SASP. The answer may lie in the very complexity of the senescent response, including SASP, and these additional responses an also offer beneficial effects such as increased immune clearance and tissue repair.

Although much of the evidence that links cellular senescence and tumor suppressor pathways comes from cell culture system, there is also a great deal of supporting evidence from intact organisms. Inactivation of the genes encoding p53 or INK4a results in cells that fail to senesce in response to several stimuli and in all the cases, these animals developed tumors at an early age [182]. Similarly, genetic manipulations in mouse models which cause premature senescence in mammary epithelial cells suppressed the development of breast cancer even after exposure to the mouse mammary tumor virus [183].

A number of studies showed that cellular senescence is induced in "premalignant" tumors, but that it is rarely induced in more advanced malignant tumors [164,179,184]. These data strongly suggest that cellular senescence is an important in vivo tumor response to prevent tumor development. Importantly, some SASP factors have been shown to induce senescence in neighboring cells. For example, the SASP component, GROα, promotes senescence in human ovarian fibroblasts [185], and PAI-1 (plasminogen activator inhibitor 1) promotes replicative senescence [186] in mouse cells. Therefore, the induction of senescence in a single cell can spread to neighboring cells which is a potent mechanism to stop cancer progression.

Besides suppressing tumorigenesis by growth arrest, senescent cells can also suppress cancer by stimulating the immune system. Senescent cells are attacked by immune cells including leukocytes, macrophages, natural killer cells and T cells of the immune system. This reaction, seems to result in the clearance of senescent cells and could also function by stimulating the local immune system to eliminate nearby oncogene-expressing cells [119]. The cytokines released in SASP that are responsible for these immune responses are not completely understood. Another function of senescence response and SASP is to

promote optimal healing of the wound after tissue injury. It has been shown that in case of acute liver injury [187] and cutaneous wounds [188] senescent cells limit the development of fibrosis.

Besides these benefits, senescent cells also contribute to age-related pathology. In many cases, senescent cells have been shown to cause degenerative changes mostly through secreted proteins found in SASP [115]. Besides causing degenerative pathology, senescent cells can also drive hyperplastic pathology. It has been thought that senescent cells accumulate with age and create a micro-environment that makes cells vulnerable to cancer development or cancer progression. This seems to be counter-intuitive as senescence prevents the proliferation of damaged cells, and they are the major risk factor for cancer development. However, senescence in response to radio- or chemo-therapeutic agents, results in the secretion of factors which have been shown to protect neighboring tumor cells from being killed by those same therapeutic agents [189,190]. There is an obvious irony in these studies of cellular senescence, but they demonstrate that the effects of senescent cells in the tumor microenvironment are quite complex and highly dependent on the physiological context.

Although senescence is not completely understood, it is clear that senescence plays a major role in tumor suppression. Importantly, many chemotherapeutic drugs have been found to induce senescence in cancer cells. A new class of chemotherapeutic agents, the histone deacetylase inhibitors, are being developed and these agents are strong inducers of senescence.

1.8.7. Using Sirtinol, a Sirtuin Inhibitor, for Inducing Senescence in Cancer Cells

In budding yeast, the *SIR* genes (silent information regulator); *SIR1, SIR2, SIR3 and SIR4* were first identified as necessary components for the transcriptional repression of the mating type loci (*HML and HMR*) [191]. Sir2 was one of the first of these genes to be characterized and it was shown to be a NAD dependent deacetylase that removes acetyl groups from another protein. It was reported in 1999, Sir2, could also increase the life span in *S. cerevisae* [192]. In this study, sir2, sir3 and sir4 were disrupted in several strains. While disruption in sir3 and sir4 caused a modest 20% decrease in life span, the sir2 mutant strains showed a 50% reduction in life span. Later studies demonstrated similar effects of Sir2 on the lifespan of worms and flies [193].

Mammals have seven sirtuins (SIRT 1-7), which have a highly conserved NAD^+-binding domain and a catalytic domain. They are categorized as class III histone deacetylases (HDACs). HDACs are a group of enzymes that catalyze the removal of acetyl groups. However, not all of them have deacetylase activity or use histones as substrates and many non-histone substrates have been characterized [194]. Sirtuins have been shown to protect against aging diseases by altering a variety of cellular mechanisms such as the stress response, DNA repair and apoptosis [195,196]. They are involved in modulating most key cellular processes such as transcriptional regulation, DNA damage repair, apoptosis, cell cycle control, metabolism and senescence [197].

The finding that sirtuins play a beneficial role in increasing life span has increased the interest in determining the effect of sirtuins in mammalian cells. Specifically Sirt1 has been extensively characterized for its role in aging [198] and age-related diseases. In fact, caloric

restriction, which has a great effect on life span [199,200] is also partly linked to sirtuin levels in several model organisms [196].

In addition, Sirt 1 has an important role in the regulation of cell fate and stress responses in mammalian cells. It promotes cell survival by inhibiting apoptosis and DNA-damage induced senescence. Several proteins have been identified as Sirt1 substrates including: p53 [201], peroxisome proliferator-activated receptor-gamma [202], FOXO transcription factors [203] and Ku 70 [204].

It is important to mention that Sirt1 has a role in cancer [205]. It has been shown that SIRT1 expression is elevated in many types of cancers. Overexpression of SIRT1 has been shown in human tumor tissues and also during experimental carcinogenesis [206,207]. This elevation indicates that SIRT1 may be a tumor promoter. SIRT1 has been suggested to have an important role in DNA damage repair, genomic integrity and inhibition of tumorigenesis [208]. SIRT1 levels are also modulated by oncogenes and tumor suppressor genes such as HIC1 and DBC1 [209,210]. Moreover, SIRT1 deacetylates non-histone proteins in cancer cells. Therefore, inhibition of SIRT1 was suggested as a potential cancer therapy and this idea is supported by studies reporting that reduced SIRT1 activity results in cancer cell death [211,212]. Similarly, the SIRT1 activator, SRT1720, has been shown to cause metastasis in a mouse model of breast cancer [213]. However, it has been shown that, increased expression of SIRT1 reduces colon cancer formation in mouse models [214]. Therefore, it is not completely clear how modulating SIRT1 results in both tumor suppression and promotion.

A number of SIRT inhibitors have been discovered and tested as anticancer agents including, nicotinamide, sirtinol, cambinol etc [215]. It has been shown that Sirt1 inhibition by sirtinol or by genetic means, results in the induction of SA-β-gal staining and the

induction of a flat and large cell morphology in cancer cells [212]. Similar results have also been seen in human diploid fibroblasts [216].

Since senescence is a potential tumor suppressor mechanism, sirtinol and other senescence inducing drugs may be useful as cancer therapeutics. In order to explore the effects of sirtinol, I used the drug to treat H1299 cancer cell line and I used a novel hybrid sample preparation method that I developed to explore changes in the proteome in response to sirtinol. This work will help to understand the long term response of cancer cells to sirtinol treatment.

2. AIMS OF THE PROJECT

For my research project, I had three major aims;

1. Developing a hybrid technique for quantitative mass spectrometry;

Obtaining and studying proteome data can be challenging in many ways. One of the biggest challenges arises from sample preparation steps. These steps can be modified according to the final goals of the study. One of the issues with mass spectrometry analysis is maximizing the number of protein identifications from the analysis. Samples where the complexity and dynamic range of the proteome is high makes these studies even more challenging. In this part of my project, I developed a sample preparation methods to increase the number of protein identifications during MS analysis.

2. Verification of the hybrid method by obtaining proteome profiles of H1299 cells after sirtinol treatments;

I developed the hybrid method using Bovine Serum Albumin (BSA) as a test protein, as it is easy to obtain and contains many cysteine residues. I also used the hybrid method to perform a proteomic analysis of H1299 cells treated with the sirtinol, an anti-cancer agent. I found that sirtinol causes a robust phenotype in H1299 cells, so it would make an ideal test compound to extend the hybrid technique to real world samples. Furthermore, by performing a pairwise comparison, I was able to demonstrate the adaptability of the hybrid method to quantitative analysis. Furthermore, working with a potentially therapeutic agent may give insights into the mode of actions of this compound.

3. Validating the proteomic analysis and comparing results with senescent mouse embryonic fibroblasts (MEFs) to obtain biologic relevance;

The final aim of my work was to verify the proteomic changes revealed by the hybrid technique using standard cell biological and biochemical means. I chose several proteins and verified the proteomic changes by western blot analysis. I included both up and down regulated proteins and I also analyzed the levels of pRb, and important mediator of senescence, whose levels appeared unchanged in the proteomic analysis. Finally, I also determined if the levels of these proteins were altered in the classical mouse embryonic fibroblast model of senescence.

3. MATERIALS AND METHODS

3.1. Reagents and Antibodies

Chemical Reagents - All chemical reagents were purchased from Sigma-Aldrich unless otherwise noted.

Standard Solutions - All solutions used in this study are described in the text with the exception of the following:

PBS buffer: 137 mM NaCl, 2.7 mM KCl, 10mM NaH_2PO_4, 2mM K_2HPO_4, 2mM K_2HPO_4, pH 7.4

Antibodies - The antibodies used in this work are listed in **Table 3.1.**

ANTIBODY NAME	COMPANY, CATALOG #
AHNAK (E-5) Antibody	Santa Cruz, sc-390743
Anti-rabbit IgG, HRP-linked Antibody	Cell Signaling, 7076
Anti-mouse IgG, HRP-linked Antibody	Cell Signaling, 7074
EIF2α Antibody	Cell Signaling, 9722
Fatty Acid Synthase (C20G5) Rabbit mAb	Cell Signaling, 3180
Pan-Actin Antibody	Cell Signaling, 4968
Phospho-Sirt 1 (Ser-47) Antibody	Cell Signaling, 2314
PPP2R3A Antibody	Abcam, 126195
p-Rb (Ser795) Antibody	Cell Signaling, 9301
p-Rb (Ser795) Antibody	Cell Signaling, 9301
p-Rb (Ser807/811) Antibody	Cell Signaling, 8516
Rb (IF8) Antibody	Cell Signaling, 9309

Sirt 1 (D1D7) Antibody	Cell Signaling, 9475
Sirt 2 (D4050) Antibody	Cell Signaling, 12650
Sirt 3 (D22A3) Antibody	Cell Signaling, 5490
Sirt 5 (D8D12) Antibody	Cell Signaling, 8782
Sirt 6 (D8D12) Antibody	Cell Signaling, 12486
Sirt 7 (D3K5A) Antibody	Cell Signaling, 5360

Table. 3.1. List of antibodies. For each antibody, names and product numbers are reported.

3.2. General Biochemical Techniques

Protein Concentration Assay

Protein concentration was measured with Bradford Reagent (Bio-Rad Protein Assay) in this study [217]. The measurements were carried out according to the manufactures instructions.

SDS-PAGE Gel Electrophoresis and Western Blot

Lysates were supplemented with 6X Laemmli sample buffer (Sigma Aldrich) and boiled for 10 minutes. The sample was separated by 12% SDS-PAGE [218] and then electroblotted onto PVDF membranes (Millipore). The membranes were blocked at room temperature for 2 hours in PBS buffer supplemented with 0.1% Tween-20 and 5% milk. The blocked membranes were then incubated with the appropriate primary antibodies, diluted in the blocking buffer and incubated with the membranes overnight at 4°C. The membranes were washed extensively with PBS supplemented with 0.1% Tween-20, and then incubated for 1 hour with a horseradish peroxidase (HRP)-conjugated secondary antibody. After extensive washing, the blots were developed with enhanced

chemiluminescence reagents (ECL) (GE Healthcare) following the manufacturer's instructions.

Colloidal Blue Coomassie G-250 Staining

To visualize proteins after SDS-PAGE, the gels were washed once with distilled water and incubated for one hour at room temperature with colloidal blue G-250 coomassie solution (10% (w/v) $(NH_4)_2SO_4$, 20% Ethanol, 0.4% (w/v) Coomassie Brilliant Blue G250 and 3% phosphoric acid.) Then gels were washed with distilled water in order to remove the background staining and to enhance the visualization of protein bands.

3.3. Cell Handling Procedures

Non-small Lung Cancer H1299 Cell Culture

H1299 cells were grown in DMEM media (Sigma Aldrich) supplemented with 10% fetal bovine serum (Hyclone), 2 mM L-glutamine, 50 U/ml penicillin, 100 ng/ml streptomycin B, and 25 ng/ml amphotericin B.

Stable Isotope Labelling of H1299 Cells

Isotopic Labeling of H1299 cells was performed with the SILAC technique [75]. SILAC medium was purchased commercially (Silantes, # 282946423) and supplemented with 10% dialyzed FBS (Silantes, #281000900) and 2mM L-glutamine. The media was also supplemented with unlabeled or labeled lysine and arginine (Lys-$^{12}C_6$ and Arg-$^{12}C_6$ or Lys-$^{13}C_6$ and Arg-$^{13}C_6$) (Silantes, #282926433). Cells were grown in SILAC medium for minimum 5 passages and mass spectrometry was used to ensure that the incorporation of heavy isotopes was complete.

Sirtinol Treatment of H1299 Cells

H1299 cells were plated at a density of 10,000-15,000 cells per cm^2 and allowed to recover overnight and were treated with the indicated concentrations of sirtinol (Adipogen, Lot No. A00055) for the indicated period. Control cells were treated with and equivalent dose of DMSO, as this served as the carrier solvent for sirtinol. After exposure to Sirtinol, the cells were washed three times with sirtinol-free medium and cultured for the indicated times in complete media without sirtinol. Cell viability was determined by the Alamar blue assay and Trypan blue exclusion test and viable cells were counted.

Alamar Blue Assay

At the appropriate time points, growth media was replaced with 500 ul of fresh media supplemented with Alamar Blue reagent (Sigma) at the final concentration of 100 ug/ml. Cells were incubated with Alamar Blue for 2 hours, then 100 ul of the media was collected, transferred to a black 96 multiwell ViewPlate™ (Perkin Elmer) and the fluorescence was measured with excitation/emission wavelengths set to 560/590 nm respectively using an EnVisionR Multilabel Plate Reader (Perkin Elmer). Values were normalized to the fluorescence obtained from adding Alamar Blue to growth media and cultured in the absence of any cells.

Mouse Primary Fibroblast Cultures

Mouse Embryonic Fibroblast (MEFs) were isolated from PTEN $^{LOX\text{-}LOX}$ mice. Pregnant females at 13-14 days of gestation were euthanized and the embryos were isolated from the uterus and washed with sterile PBS three times. Heads and livers were removed and each embryo was minced on a plate with a scalpel. Each diced embryo was transferred to

its own tube and 500 ul of pre-warmed trypsin-EDTA solution (Sigma-Aldrich) was added to the tubes and incubated at 37°C for 10 minutes. The tissue was then homogenized by trituration through pipette tips. The content of each tube was transferred to one 25 cm^2 flasks with 5 ml of complete media (DMEM media (Sigma Aldrich) supplemented with 10% fetal bovine serum (Hyclone), 2 mM L-glutamine, 50 U/ml penicillin, 100 ng/ml streptomycin B, and 25 ng/ml amphotericin B) and the MEFs were left in culture with periodic media changes until they reached full confluence. The MEFs were then maintained on the 3T3 protocol [166].

SA-βgal Staining

Cultured H1299 cells were washed with PBS and fixed with 3.7 % formaldehyde for 5 minutes and incubated overnight at 37 °C in freshly prepared X-gal staining buffer (1mg mL^{-1} X-gal (5-bromo-4chloro-3-indolyl β-D-galactosidase), 5 mM $K_3Fe[CN]_6$, 5 mM $K_4Fe[CN]_6$ and 2 mM $MgCl_2$ in PBS, pH 6.0. At the end of the incubation period, the cells were washed with H_2O and examined at x20 magnification.

Cell Lysis

Cells were washed twice with PBS and then lysed in ice-cold NP40 buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% NP40), and the lysate sonicated in 0.5 second bursts at 50% power for 1 minute using a vial tweeter (Hielscher Ultrasound Technology), and spun at 13,000 rpm for 10 minutes to remove cell debris.

3.4. General Proteomics Sample Preparation Methods

In-solution Digestion

Dithiothreitol (DTT) was added to a final concentration of 5 mM and the sample was heated to 55°C for 10 minutes. Iodoacetamide or Chloroacetamide was added to 15 mM and the reaction was allowed to proceed for 1 hour at room temperature. Trypsin was then added at a ratio of ~1:20 (Trypsin: substrate) followed by digestion for overnight at 37°C or 1 h at 57°C. The digested sample is desalted and concentrated using filter tips [219].

In-gel Digestion

Samples separated by SDS-PAGE were stained with Commassie blue as previously described. Visualized bands, or discrete areas of the gel, were excised using a scalpel and chopped into ~1mm^3 pieces. Reduction and alkylation steps are performed similarly to the in solution digest protocol and then subjected to extensive washing with 20 mM Triethyl Ammonium Bicarbonate, pH 8.5 in 50% Acetonitrile. The gel pieces were dehydrated with 100% Acetonitrile and then 20 ng/ul trypsin (Promega) was added to completely cover the gel pieces and the digestion was allowed to proceed overnight at 37°C. The supernatant was harvested and any remaining peptides were extracted from the gel by sonicating for 20 minutes with 1% Formic acid (FA). Recovered tryptic peptides are then purified using filter tips as described.

3.5. Hybrid Method

Performic Acid Treatment

Performic acid was produced by mixing equal concentrations of formic acid and hydrogen peroxide. The mixture was left to react at room temperature for 1h before use. Performic acid was added to a final percentage of 3% in the sample to be oxidized.

Preparing PolyWAX in-Tip Column

PolyWAX resin (PolyLC) was washed with 30 bed volumes of Acetonitrile (Fluka) and then with 30 bed volumes of HPLC-grade Water (Fluka). The resin was then incubated with a concentrated salt solution for at least 12 hours. Prior to its initial use, the resin is washed with 30 bed volumes of HPLC-grade water to remove the salts and finally equilibrated with 30 bed volumes of binding buffer (e.g., 70% ACN, 10 mM Ammonium Bicarbonate, pH 7).

Preparing Stage Tips

Stage tips were prepared with reverse phase filters (Empore, SDB-XC, 710265). The volume of the filter disk and the tips are determined based on the amount of sample used. Reverse phase filters are used both for sample desalting and as a frit in cysteic acid enrichments steps. A small core of the reverse phase filter is punched out of the filter using a 2.5 ml combi-tip (Eppendorf) and then lodged near the end of a 20 or 200 ul pipette tip (Eppendorf).

For enrichment of cysteic acid containing peptides, PolyWAX resin (Poly LC, 10-139-2) is used. The indicated amount of conditioned PolyWAX resin was layered on top of the

reverse phase filters. Loading, washing and elution steps are performed by pushing the appropriate solvent over the tip.

The Hybrid method is used both with in-solution and in-gel solution approaches.

Hybrid Method with in-Solution Digestion

For BSA samples, various amounts of BSA samples were digested with sequencing grade-modified trypsin (Promega) overnight at 37°C or 1 h at 57°C. Samples were oxidized with performic acid (3%) for 1h at room temperature and desalted with 0.1% FA using filter tips (Empore SDB-XC). The tryptic peptides are then fractioned off-line with ERLIC chromatography using weak anion exchange resin (Poly LC, PolyWAX LP^{TM}) in a filter tip. The volume of the resin was adjusted according to the amount of sample. Mobile phase was 70% ACN and 20 mM Ammonium Bicarbonate at pH 7 and cysteic acid containing peptides were eluted with the indicated buffer. All fractions of in-tip chromatography were analyzed by LC-MS/MS.

Hybrid Method with in-Gel Digestion

For complex samples from tissue culture cells, 25 ug protein was separated by SDS-PAGE and the gels were stained with Commassie blue as described. Instead of reduction and alkylation steps, cysteine residues were oxidized by addition of performic acid (3%) to the excised gel pieces for 1 h at room temperature. The gel pieces were washed extensively and digested with sequencing grade-modified trypsin (Promega) for overnight at 37°C, as previously described. Peptides were subsequently extracted from the gel with 5 % FA. Recovered tryptic peptides were then fractionated off-line with ERLIC chromatography using PolyWAX resin, as previously described. Finally, the peptides were analyzed with LC-MS/MS.

Making Methyl Esters

Methanolic HCL reagent was made by adding 800 ul acetyl chloride slowly to 5 mL of dry methanol while stirring. 50 ul methanolic HLC was added to desalted and lyophilized tryptic peptides. The mixture is let stand for 2 hours at room temperature.

3.6. LC/MS-MS Analysis

Chromatographic separation was accomplished using 3 um C12 resin (Phenomenex) packed into a 75 um x 15 cm fused silica capillaries (Polymicro) using an EASY-nLC II system (Bruker) at a flow rate of 500nl/ml. 90 minute gradients from buffer A (99.9% water with 0,1% FA) to buffer B (100% methanol) were used develop the column. The column effluent was sprayed directly into an Amazon ETD mass spectrometer (Bruker Daltonics) using a homemade electrospray interface.

Data collection was controlled by the HyStar program (Bruker) using a data-dependent acquisition mode. MS spectra were acquired in the range of *m/z* 380–1600 followed by maximum of five MS/MS analyses. The typical duty cycle was approximately 2 seconds. The precursor selection was based on peak intensity and all charge states were selected except for 1+, which was excluded. Dynamic exclusion was set to allow each peak to be analyzed 3 times in a 1 minute window.

3.7. Data Analysis

X!Tandem

Data analysis was performed on 3 independent replicates for each experimental condition. Data from each experiment was merged and the searches were performed on both merged data and on individual experimental runs using X!Tandem. Data was

searched against the UniProt human database (forward and decoy). Trypsin was selected as protease with maximum one missed cleavage allowed. Carbomidomethyl or full oxidation of cysteine (47.97-C) were set as complete modifications and partial oxidation of methionine and partial deamidation of glutamine and asparagines were also allowed in the search parameter. For SILAC labeled samples, $^{13}C_6$ arginine & $^{13}C_6$ lysine are chosen as partial modifications. MS tolerance was set to 500 ppm while MS/MS tolerance was set to 0.4 Da.

Mascot

Data analysis was performed on 3 independent sample replicates for each experimental condition. Data from each experiment were searched individually on Mascot. Data was searched against the UniProt human database (forward and decoy). Trypsin was selected as protease with maximum one missed cleavage allowed. Full oxidation of cysteine (47.97-C) is chosen for complete modifications. Deamidation (NQ), oxidation (M), are chosen as partial modifications. Peptide tolerance was set to 0.5 Da. MS/MS tolerance is set to 0.5 Da. SILAC support is built into MASCOT and SILAC was chosen as quantification method when needed.

3.8. Peptide/Protein Verifications and Quantifications

ProteoIQ software was used for verifying peptide identifications and also for SILAC based quantifications. Peptide probability and peptide false discovery rate were used for establishing a threshold of confidence for peptide identifications generated from the database search tools. Peptide probability is set to 0.2 (80% likelihood) and the protein probability is set to 0.5 (50% likelihood). False discovery rate (FDR) was estimated based on decoy database searches in ProteoIQ software. FDR cut-off was set to 5%.

The quantification based on SILAC is performed with ProteoIQ Software. For every peptide, the precursor ions are detected based upon the SILAC parameters specified. Precursor ion intensities are then saved for each peptide throughout the scan. Ratios are calculated by dividing each precursor ion intensity by the intensity of the reference group precursor ion.

4. RESULTS

In mass spectrometry based proteomics, sample preparation is critically important because it has an impact on the quality of the data produced and in the overall number of protein identifications. There are many methods in proteomics that have been developed for different purposes.

I developed a sample preparation technique, which selectively enriches the oxidized cysteines and cystines in proteomic mixtures. BSA was used as a target protein for the development and optimization of the technique. However, more complex samples such as cell lysates were also used for further improvement of technique. Several factors including the appropriate reagents, the pH and the column capacity were found important in the efficiency of the hybrid technique. Results show that the hybrid method greatly increases the number of peptide and protein identifications when compared to classical methods, Besides providing increased numbers of identifications, the hybrid technique was also efficient in increasing the identification of several different PTMs such as phosphorylation, acetylation etc. The ease of implementation and the reproducibility were also found to be advantages of the technique.

In the second part of the project, I chemically induced cancer cells to enter senescence with sirtinol, a Sirt1 inhibitor. I used the hybrid technique that I developed to investigate the effect of sirtinol on the proteome of the treated cells. I used a variety of approaches to verify the changes in the proteome identified using my technique. Mass spectrometry analysis revealed that 220 proteins were differentially expressed, which indicates the utility of the hybrid technique in binary comparisons. Some of these proteins have already been investigated for their role in senescence and cancer, while the role of other candidate proteins in senescence is currently unclear. Moreover, bioinformatic software

revealed that the acetylation pathway was strongly affected after sirtinol treatment which is consistent with the nature of the drug.

4.1. Development of the Hybrid Technique

I developed a protocol taking advantage of the conversion of cysteine to cysteic acid by peroxide treatment. The method can be coupled to SILAC for accurate quantification of proteins. As a result, my method has the ease of labeling and enhanced quantitative performance of SILAC protocols with that added benefit of the reduction in sample complexity seen with ICAT protocols and my method can be thought of as a hybrid between these two techniques.

I chose to perform the initial optimizations using BSA as a target protein. BSA is cheap, abundant and contains many cysteine and cystine residues, which makes it particularly useful for optimizing the steps involving the enrichment of cysteic acid containing peptides. Although, the idea of the hybrid technique is conceptually simple, there were many steps that required optimization.

The method has mainly two steps; creating cysteic acid residues by oxidation of the cysteine residues in peptides and enrichment of these peptides with an optimized form of chromatography.

4.1.1. Creating Cysteic Acid Residues

The cysteine thiol groups in peptides are nucleophilic and can be easily oxidized. However the level of oxidation can vary depending on the oxidants. The most common reaction of this groups is a reversible oxidation, which forms a disulfide bond between two molecules of cysteine to cystine.

Complete oxidation of cysteine occurs in three steps. In the first oxidation step, the hydrogen on the sulfur atom is replaced with a single hydroxyl group (OH) to create cysteine sulfenic acid. A second oxidation steps adds an oxygen atom to create cysteine sulfinic acid. The third and final oxidation step adds another oxygen atom to create cysteine sulfonic acid, which is also called cysteic acid **(Figure 4.1)** [220]. For the analysis of oxidized cysteine residues by mass spectrometry, it is important that the cysteines be fully oxidized, as incomplete oxidation will make quantification difficult and unnecessarily increase the sample complexity.

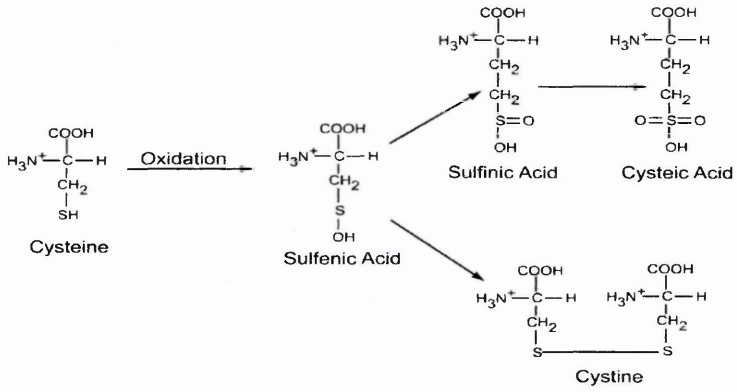


Figure 4.1. Conversion of cysteine to cysteic acid by oxidation [220]**.** Oxidation of cysteines schematized in three steps. In the first oxidation step, the hydrogen on the sulfur atom is replaced with a hydroxyl group (OH), creating sulfenic acid (cysteine sulfenic acid). The second oxidation adds another oxygen atom to produce a sulfinic acid (cysteine sulfinic acid). A third oxidation step adds another oxygen atom to produce a sulfonic acid (cysteic acid). Cystine is the amino acid formed by the oxidation of two cysteine molecules to form a covalent link via a disulfide bond.

In my initial experiments, I tried several different oxidants for oxidizing the cysteines and cystines. These studies involved the oxidation of BSA and also complex samples such as H1299 cell lysates with different oxidizing agents and the use of different agents to halt the oxidation. For example, using hydrogen peroxide (even at high concentrations) was not enough to completely oxidize cysteines and cystines and other very strong oxidizing agents, such as perchlorate, also resulted in the destruction of the sample. I have found

that performic acid (a mixture of formic acid and hydrogen peroxide) is a quick and efficient reagent for oxidizing cysteines and cystines without damaging non-sulfur containing amino acids. While hydrogen peroxide was able to oxidize approximately 35 % of the cysteine containing peptides, performic acid was able to oxidize all the cysteine containing peptides in the sample **(Figure 4.2)**.

Treatment time and concentration of performic acid were also optimized using several different combinations and a 1 hour treatment with a final concentration of 3% performic acid was found be optimal.

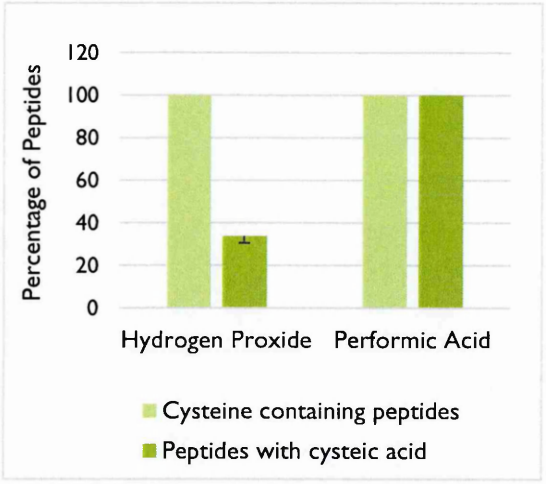


Figure 4.2. Comparison of hydrogen peroxide and performic acid in cysteine oxidation. Total cell lysates were separated by SDS-PAGE gel and gel pieces were treated either with hydrogen peroxide (200 uM, 2 h, room temperature) or performic acid (3%, 2h, room temperature). After peptides were digested and extracted from gel pieces, LC-MS/MS experiments were performed. The percentage of oxidized cysteines were calculated for each sample. The total number of cysteine containing peptides and the total number of cysteic acid containing peptides is shown. Assays were performed in triplicate and the graph displays the mean +/- s.d.

In addition, I examined where during the protocol it is best to perform the performic acid treatment. The treatment step can be quite critical as any extra steps required to prepare the sample for treatment, or to remove the peroxides can greatly impact on the overall efficiency of the method. I found that performic acid treatment can be used directly in

place of the classically used reduction and alkylation steps and therefore no additional steps were needed. Whereas, treating the samples with performic acid prior to gel electrophoresis requires extra steps. After the cysteines and cystines are oxidized to cysteic acid, I began searching for a technique to specifically enrich these peptides.

4.1.2. Specific in-Tip Chromatography Technique for Cysteine Enrichment

The use of several different chromatographic techniques were explored that in principle can be used for the selective enrichment of cysteic acid containing peptides such as strong cation exchange chromatography (SCX), strong anion exchange chromatography (SAX) and electrostatic repulsion and hydrophilic interaction chromatography (ERLIC). These chromatography techniques are different variants of ion exchange chromatography.

These chromatography techniques are usually performed with dedicated chromatography equipment and the columns are developed with gradients of different buffers. However in this study I took advantage of using these systems off-line, in-tip. This in-tip technique has been previously described as Stop and Go Extraction Tips (Stage-Tips) where small disks of reverse phase beads embedded in a Teflon meshwork is placed into pipet tip as a microcolumn. However the embedded beads can also be used as a porous frit [221]. This technique has been altered to be used in enrichment of cysteic acid peptides. In this case, C_{18} disks are used as a frit and different types of resins (SCX, SAX or ERLIC) are packed on top of C_{18} disks. Loading, washing and enrichment steps are done off-line with these tips **(Figure 4.3.)**. All of the fractions are then analyzed with mass spectrometer.

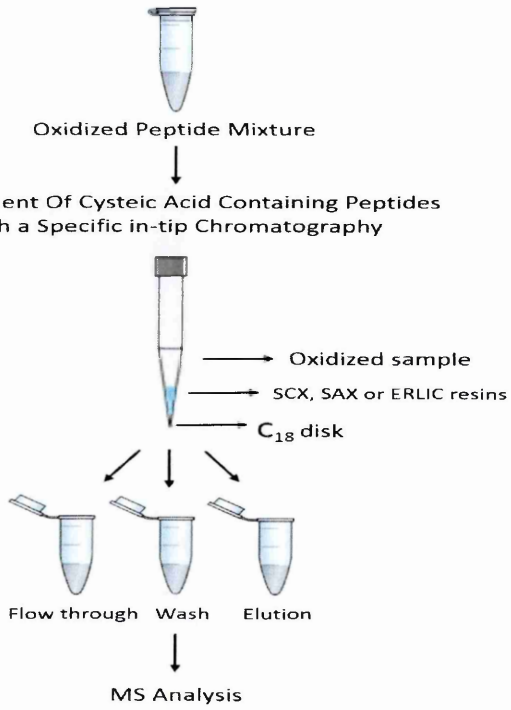


Figure 4.3. General workflow for cysteic acid enrichment experiments. The enrichment process was performed with the help of pipet tips. Depending on the experiment, SCX, SAX or ERLIC resins were used. C_{18} disks, which is a reverse phase resin embedded into a filter, were used to form a frit. Oxidized peptide mixtures were passed through pipet tip and collected (flow through). The mixture was then washed with binding buffer to remove unspecific bindings (wash) and finally all the peptides left in the column were removed with elution buffer (elution) and all fractions were analyzed by mass spectrometry.

4.1.2.1. SCX for Enrichment of Cysteic Acid Containing Peptides

I initially attempted to use Strong Cation Exchange (SCX) chromatography for cysteic acid enrichment. In the case of SCX, the cysteic acid containing peptides were expected to elute very early from the column as, SCX binds to positively charged peptides. The early elution of negatively charged residues has been used extensively in phospho-proteomics. However, the cysteic acid containing peptides eluted from the column in each step and were not enriched in the flow through. The difference in behavior between phospho-peptides and cysteic acid containing peptides, is most likely caused by the

extremely low pKa of the phosphoryl group. I have tried different binding and washing conditions, especially in the range where cysteic acid will be ionized, the results indicated that SCX was failing to enrich for cysteic acid. Some of the best results were found at pH 7 and these are shown in **Figure 4.4.**

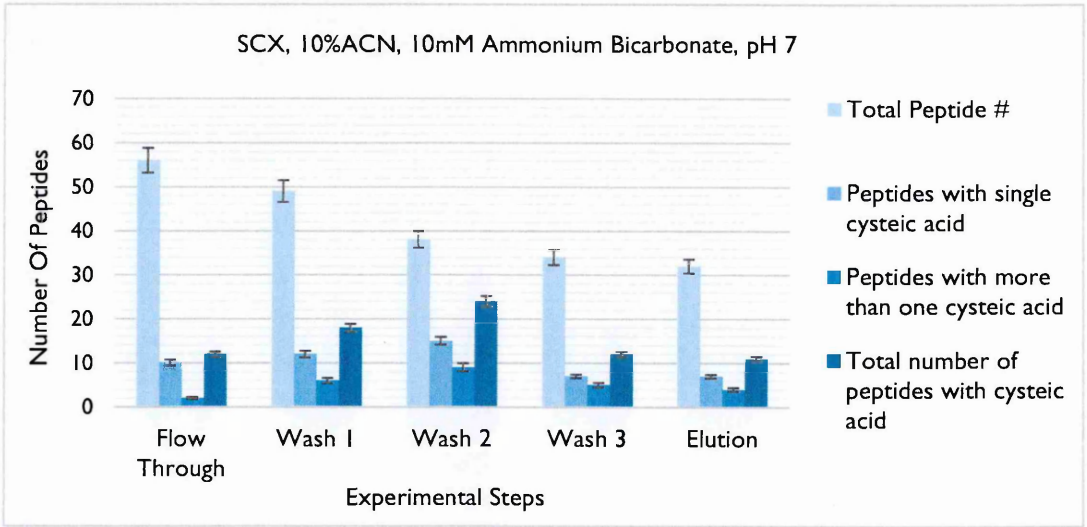


Figure 4.4. Enrichment of cysteic acid peptides with SCX. BSA was oxidized and digested and the resulting peptides were loaded onto SCX columns and washed with 10% ACN, 10 mM ammonium bicarbonate at pH 7. Peptides were eluted from the column with 70 % ACN and 0.1 % TFA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graph shows the number of peptides (total number of peptides, number of peptides with single cysteic acid, number of peptides with more than one cysteic acid and the total number of cysteic acid containing peptides) that were identified in each experimental step (flow through, washes and elution). (+/- s.d.)

4.1.2.2. SAX for Enrichment of Cysteic Acid Containing Peptides

After the failure of SCX, I decided to use SAX columns for cysteic acid enrichment. SAX works oppositely to SCX, and in SAX chromatography the negatively charged peptides bind strongly to the column, while the positively charged peptides are eluted from the column in earlier steps. In a manner similar to SCX, there was no significant enrichment of cysteic acid containing peptides **(Figure 4.5.).**

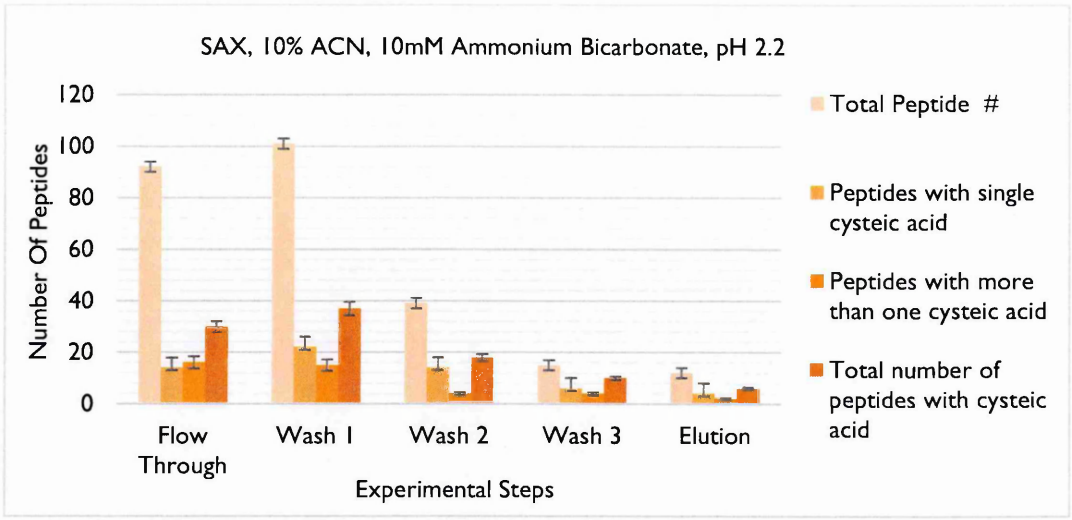


Figure 4.5. Enrichment of cysteic acid peptides with SAX. BSA was oxidized and digested and the resulting peptides were loaded onto SAX columns and washed with 10% ACN, 10 mM ammonium bicarbonate at pH 2.2. Peptides were eluted from the column with 70 % ACN and 0.1 % TFA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graph shows the number of peptides (total number of peptides, number of peptides with single cysteic acid, number of peptides with more than one cysteic acid and the total number of cysteic acid containing peptides) that were identified in each experimental step (flow through, washes and elution). (+/- s.d.)

After I analyzed the list of the peptides eluted from each step, I realized that background from peptides with Aspartic acid (D) or Glutamic acid (E) were quite high and the failure to enrich for cysteic acid may be due to these residues competing for binding on the column. In order address this, I decided to neutralize these residues by the formation of O-methyl esters. Methyl esterification was performed after cysteine oxidation. Although it helped reducing the binding of Aspartate and Glutamate to the SAX resin **(Figure 4.6)**, the overall coverage of cysteic acid containing peptides did not change. The failure of both SAX and SCX suggests that additional characteristics of the peptides are strongly influencing their behavior on classical ion exchange resins.

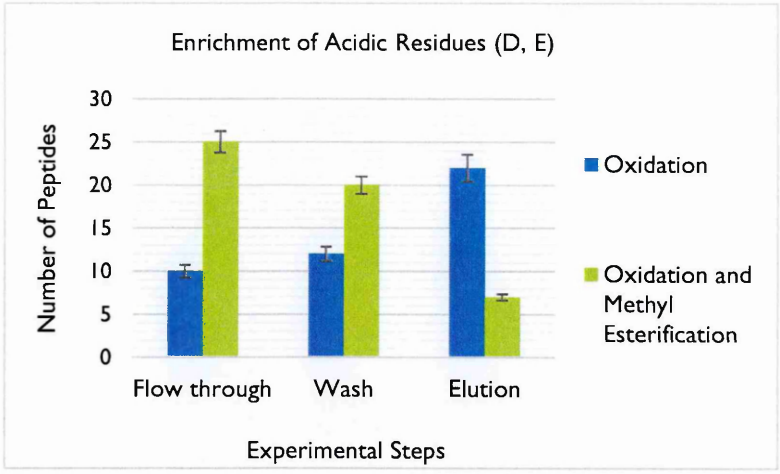


Figure 4.6. The effect of methyl esterification for the enrichment of acidic residues (Asparagine (D) and Glutamine (E)) in samples. BSA was oxidized and digested and split into equal pools. One pool was methyl esterified to neutralize the acidic amino acids and the other pool was treated similarly except that no Acetyl chloride was used to induce esterification. Pools were loaded onto conditioned SAX columns and washed with 10% ACN, 10 mM ammonium bicarbonate at pH 2.2. Peptides were eluted from the column with 70 % ACN and 0.1 % TFA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graph shows the number of enriched acidic residues in each fraction (flow through, wash and elution) for each treatment. Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.1.2.3. ERLIC for Enrichment of Cysteic Acid Containing Peptides

After these experiments, I decided to use a mixed mode chromatography technique, electrostatic repulsion and hydrophilic interaction chromatography (ERLIC) [30]. ERLIC uses a combination of the hydrophilicity and charge of the peptides for separation. It is a specific mode of chromatography that allows many separations to be done isocratically that normally require a gradient. In ERLIC, the mobile phase contains enough organic solvent so that hydrophilic interaction keeps the solutes on the column despite electrostatic repulsion. It has been used specifically for phospho-proteomic studies. Since phosphates add negative charges to the column, I hoped that ERLIC may also work for enrichment of cysteic acid containing peptides.

Like in the other tested chromatography methods, ERLIC was used off-line, in-tip, for cysteic acid enrichment [215]. C_{18} disks are used to form a frit and PolyWAX resin is packed on the top of C_{18} disks. Loading, washing and enrichment steps are done off-line with these tips.

Parameters for ERLIC chromatography are different than ion-exchange chromatography. Normally ERLIC uses high percentage of organic phase and low pH and an appropriate buffer for example, 65% ACN 20mM NH_4-formate, pH 2.2. In this mode, I aimed to keep the cysteic-acid containing peptides on the column and elute them at last with no gradient used for the mobile phase.

Starting with a high concentration of organic solvent (70% ACN) at pH 2.2., the enrichment of cysteic-acid-containing peptides was already improved relative to SCX and SAX **(Figure 4.7)**. Although the enrichment in the eluate was clearly visible some cysteic-acid containing peptides were lost in the binding and washing steps. I therefore began to explore ways to further optimize the binding of cysteic acid peptides to the PolyWAX resin.

I believed that there may be some residual salt contamination that would prevent some cysteic acid containing peptides from binding to the PolyWAX resin. Therefore, I added more washes to the in-gel digestion protocol and I switched to a buffer that was more volatile than Ammonium Bicarbonate. In this case, Triethyl Ammonium Bicarbonate was chosen as it is highly volatile and is at the correct pH. The effect of this change was unconvincing and I decided to optimize the pH of the binding step. Importantly, the ERLIC separation is highly dependent on the charge state of the peptides with both attractive and repulsive forces playing a role in selectivity.

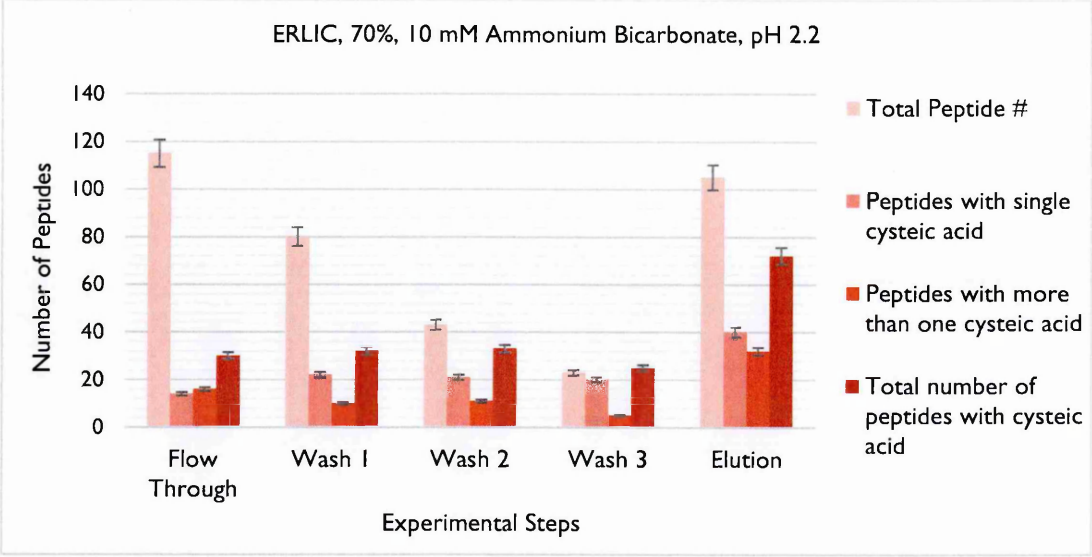


Figure 4.7. Enrichment of cysteic acid peptides with ERLIC. BSA samples were oxidized and digested. Sample mixtures were loaded to conditioned ERLIC column and washed with 70% ACN, 10 mM ammonium bicarbonate at pH 2.2. Peptides were eluted from the column with 70 % ACN and 0.5 % FA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graph shows the number of peptides (total number of peptides, number of peptides with single cysteic acid, number of peptides with more than one cysteic acid and total number of cysteic acid) that are identified in each experimental step (flow through, washes and elution). (+/- s.d.)

4.1.2.4. Effect of pH on Cysteine Enrichment

pH is one of the most important factor that affects retention on ion exchange resins, as the pH can change the net charge state of the peptides. For example, SCX is normally done at low pH to minimize the repulsion due to negatively charged residues, such as Aspartate. Normally in ERLIC, the pH of the mobile phase is selected to make sure that the solutes have the same charge as the column. For example, when phosphopeptides are separated with ERLIC, the chromatography is normally performed at a pH of 2. This is low enough to neutralize the carboxyl groups in peptides. However, the phosphate groups retain their negative charge and are attracted to the column electrostatically.

The selectivity and the capacity of a weak ion exchanger is different at different pH values. In order to see the effect of pH on enrichment, I repeated the experiments with a range of pH (from 2.2 to 8.5) to find the pH that gives the best separation of cysteic acid. Importantly, pH 7 produced the highest enrichment of cysteic acid containing peptides **(Figure 4.8.)**. Significantly, even at the optimal pH, the selectivity can be lost when the columns are overloaded.

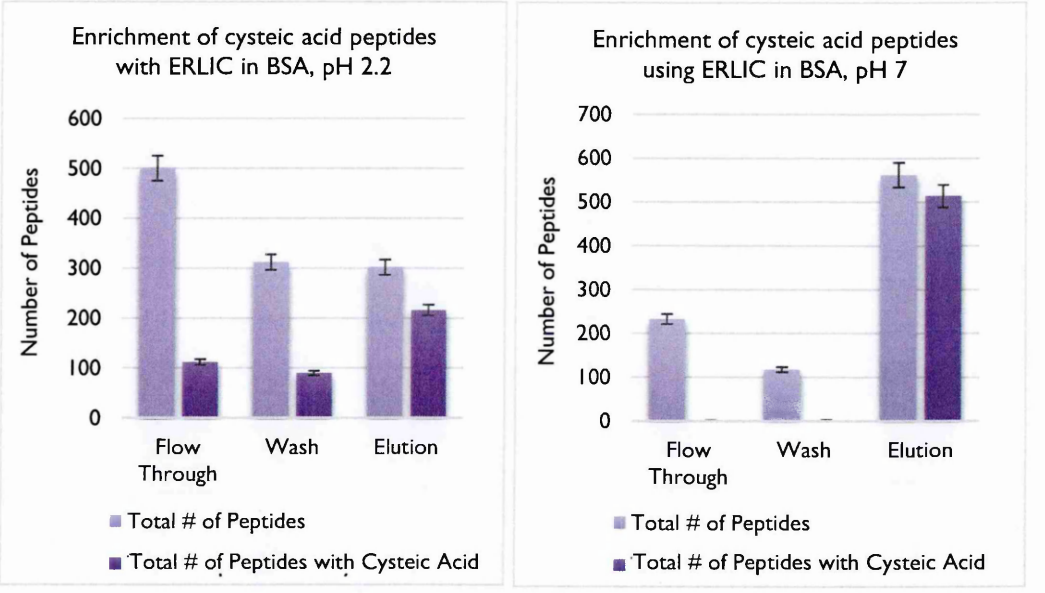


Figure 4.8. The effect of pH on the enrichment of cysteic acid peptides. BSA was oxidized and digested and loaded onto ERLIC columns and washed with 70% ACN, 10 mM ammonium bicarbonate at several different pH values (only pH 2.2 and pH 7 are shown). Peptides were eluted from the column with 70 % ACN and 0.5 % FA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graph shows the number of peptides (total number of peptides and total number of cysteic acid containing peptides) that were identified in each experimental step (flow through, wash and elution). Assays were performed in triplicate and the graphs display the mean +/- and s.d.

4.1.2.5. Column Volume and Loading Amount Effects

Column size often effects its performance, and this is especially true when the volume of resin has been reduced to enhance the overall sensitivity of the method. However, when the column begins to reach its capacity, it also begins losing its ability to separate all the

components, because strongly binding components displace weaker binding components. Therefore, I determined the maximum amount of sample that can be loaded on a certain amount of resin without losing the ability to retain cysteic acid containing peptides. I initially tested this by evaluating the performance of identical ERLIC columns to enrich different amounts of BSA and doubling the amount of BSA loaded onto the column, resulted in a breakdown in separation performance **(Figure 4.9.)**.

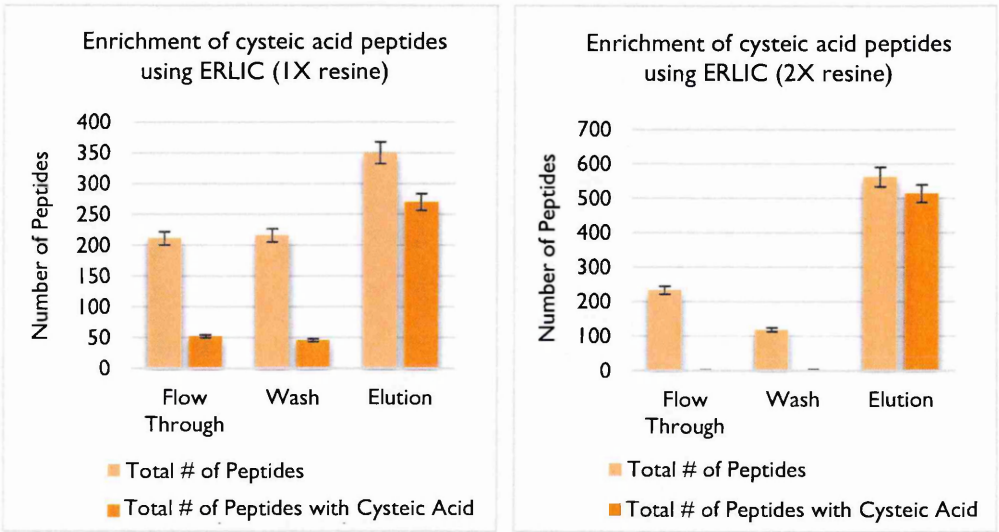


Figure 4.9. The effect of column volume on the enrichment of the cysteic acid peptides. Different amounts of BSA (1X BSA and 2X BSA) were oxidized and digested. Sample mixtures were loaded onto identical ERLIC columns and washed with 70% ACN, 10 mM ammonium bicarbonate at pH 7. Peptides were eluted from the column with 70 % ACN and 0.5 % FA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graph shows the number of peptides (total number of peptides and total number of cysteic acid containing peptides) that were identified in each experimental step (flow through, wash, and elution). Assays were performed in triplicate and the graphs display the mean +/- s.d.

However, in practice it is probably easier to scale up the column size, rather than adjusting the amount being loaded onto the column. The effect of doubling the amount of ERLIC resin in the column is shown **(Figure 4.10)**, which demonstrates that the column size can be effectively adjusted to handle more sample.

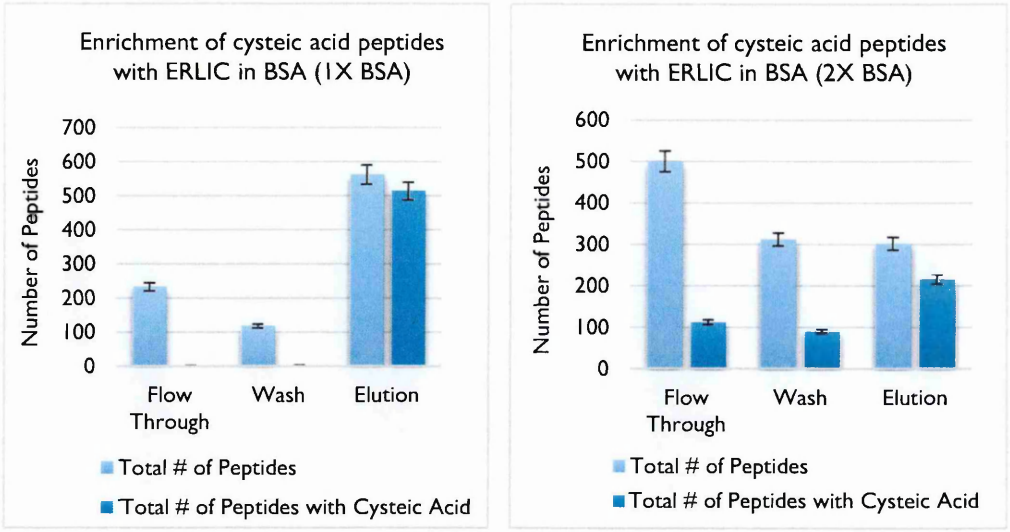


Figure 4.10. The effect of loading amount on the enrichment of the cysteic acid peptides. Different amount of BSA samples (1X BSA and 2X BSA) were oxidized and digested. Sample mixtures were loaded to ERLIC column and washed with 70% ACN, 10 mM ammonium bicarbonate at pH 7. Peptides were eluted from the column with 70 % ACN and 0.5 % FA. After enrichment, peptides from all fractions were analyzed with LC-MS/MS. The graphic shows the number of peptides (total number of peptides and total number of cysteic acid containing peptides) that are identified in each experimental step (flow through, washes, and elution). Assays were performed in triplicate and the graphs display the mean +/- s.d.

4.1.2.3. Reducing Unnecessary Steps

In the initial optimization of the ERLIC chromatography at least 3 wash steps were employed with the idea that these wash steps would improve the enrichment of cysteic acid containing peptides and to remove non-specific binding. The wash buffer was same as loading buffer (70% ACN, 10 mM AB, pH 7). However, since every fraction (loading, washes and elution) is analyzed in hybrid method, including all three of the wash steps, these extra steps increase the total time amount of analysis time. This effect is even bigger when working with samples that require multiple repeats. More importantly, I realized that the extensive washing starts removing cysteic acid containing peptides, because the large volume of wash buffer begins to develop the column in isocratic gradient mode. In order to overcome these two problems, I tested the effect of reducing the wash steps to

one. The separation efficiency of this shortened protocol can be easily seen in Figures 4.9 and 4.10. Eliminating the wash steps decreased the overall steps of the method and reduced the total analysis time **(Figure 4.11)**.

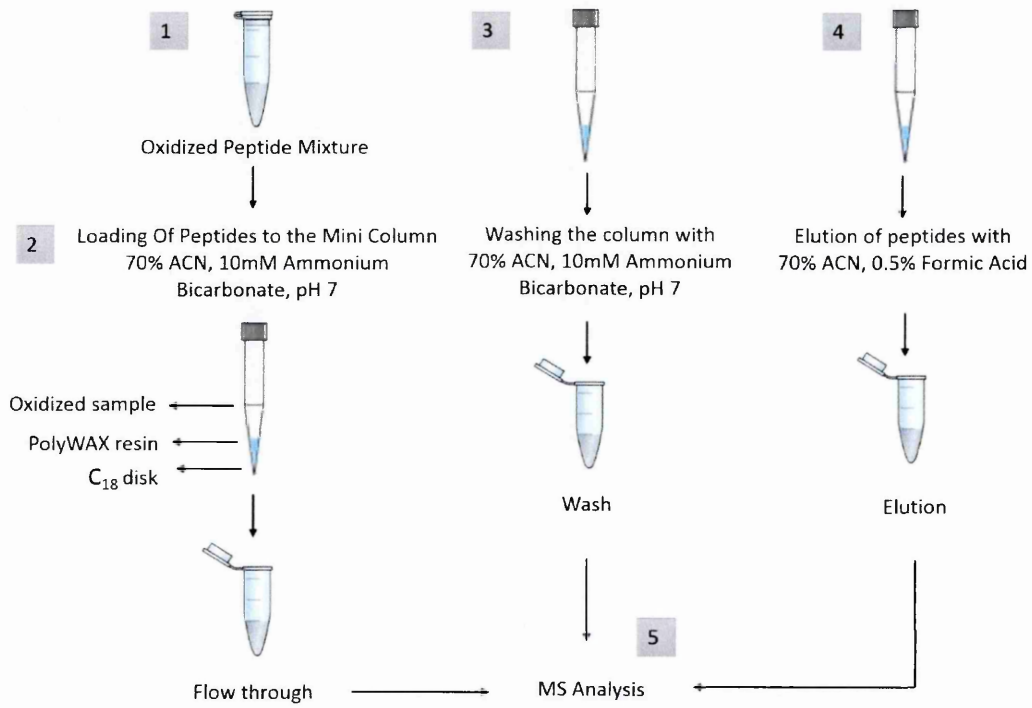


Figure 4.11. Workflow of cysteic acid peptide enrichment with ERLIC chromatography. Enrichment of cyteic acid peptides is performed in five steps. **1.** Sample is oxidized. **2.** Oxidized sample is loaded to conditioned tip prepared with PolyWAX resin and C_{18} disk. Sample is passed through pipet tip with a help of syringe and flow through is collected. **3.** Sample is washed with loading buffer (70 % ACN, 10 mM Ammonium Bicarbonate, pH 7) and wash is collected. **4.** All the peptides left in the colmun are eluted with elution buffer (70% ACN, 0.5 % FA). **5.** Each collected samples are analyzed by mass spectrometry.

4.1.3. Using Hybrid Technique in BSA Analysis

4.1.3.1. Efficient Enrichment of Cysteine Containing Peptides

At this point, the hybrid technique was quite efficient in enrichment of cysteic acid containing peptides. There were only a few cysteic-acid containing peptides in the flow

through and wash steps coming from BSA samples. Furthermore, ~90% of the peptides in eluate contained at least one cysteic acid **(Figure 4.12),** which demonstrates that the separation is both efficient and specific. The final optimized conditions for the in-tip ERLIC chromatography were 70% of ACN, 10mM Ammonium bicarbonate, at pH 7 for loading and washing steps) and the remain peptides were eluted with 70% of ACN, 0.1% FA.

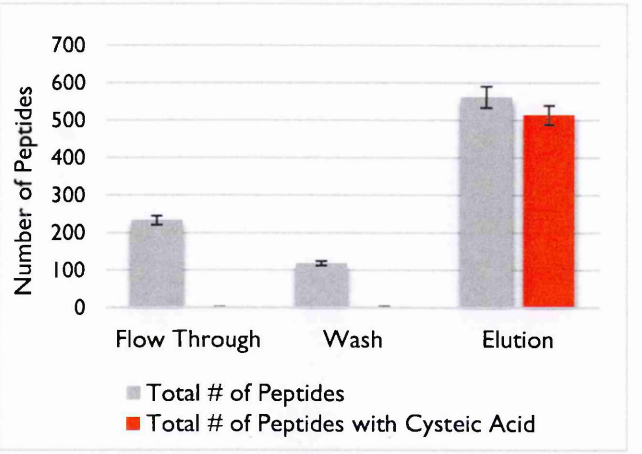


Figure 4.12. Efficient enrichment of cysteic acid peptides using ERLIC. BSA was oxidized and digested and the cysteic acid containing peptides were enriched using ERLIC chromatography. The graph shows the number of peptides (total number of peptides and total number of cysteic acid containing peptides) that are identified in each experimental step (flow through, washes, and elution). Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.1.3.2. Peptide Overlap between BSA Fractions

One measure of the separation efficiency of a new method is to determine the number of shared components between fractions. Ideally, the fractions should contain few shared components. However, even if there are many shared components, this information can be leveraged to make the method faster. For example, if there is a great deal of overlap between the flow through and wash fraction, then these fractions can be pooled together and analyzed by MS as one fraction. Therefore, I manually curated the optimized ERLIC runs to determine the overlap of peptides between fractions. There was a 15-20% overlap

of peptides in flow through and wash steps and there were no overlapping peptides shared with either of these two fractions and the elution step **(Figure 4.13.)**. This shows that ERLIC is a powerful method for peptide separation and that there is no need to pool the flow through and wash steps, as even here, at least 80% of the peptides are unique.

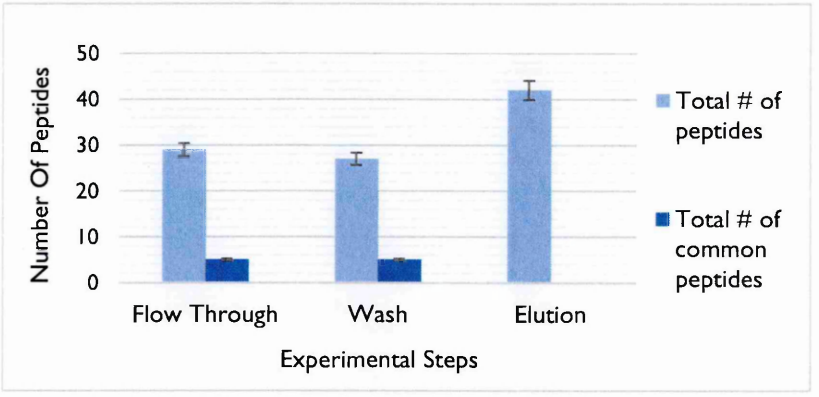


Figure 4.13. Total and common peptides that are identified in BSA samples. Enrichment of cysteic acid peptides was performed as previously described and the common peptides found between three experimental steps were quantified. The total number of peptides and number of common peptides between experimental steps are shown. Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.1.4. Using the Hybrid Technique with Complex Samples

Although, the results with BSA are very promising, it is important to demonstrate that the hybrid method will also work on more complex, "real life" samples, such as tumor cell lysates. Cell lysates are incredibly complex samples and contain a multitude of protein isoforms at widely varying concentrations. As I have shown, the hybrid method is very streamlined, so I decided to help improve the analytic power of my method by prefractionating the lysates by SDS-PAGE. The cell lysates are much more complex than a single protein, and the overall sample complexity was reduced by separating the proteins into 10 bands. The gel bands were processed and purified using the in-gel hybrid method as previously described.

As expected, enriching cysteine-containing peptides from the whole cell lysates has proven more difficult than purifying them from BSA. This is likely due to the increased sequence complexity of cell lysates and the fact that the ratio of cysteine containing peptides is much lower. In order to fix this problem, I tripled the amount of resin that I normally use with the same amount of BSA in order to increase the selectivity of the column for cysteic acid containing peptides **(Figure 4.14)**.

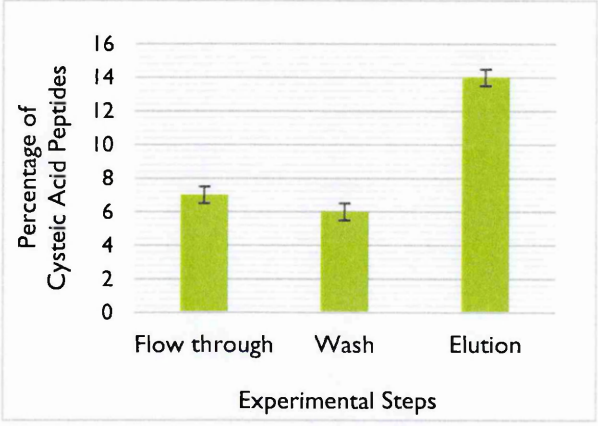


Figure 4.14. Enrichment of cysteic acid peptides by ERLIC from complex mixtures. H1229 total cell lysates were separated by SDS-PAGE for 30 minutes and the whole lane was divided into 10 equal sized bands. Each band was subjected to the hybrid purification protocol as previously described. All fractions (flow through, wash and elution) were analyzed with LC-MS/MS and the enrichment of cysteic acid containing peptides was determined and is expressed as the percentage of cysteine containing peptides. Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.1.4.1. Increased Protein Identification by Hybrid Technique in Complex Samples

The main purpose for developing the hybrid method was to increase the number of protein identifications from mass spectrometry based analysis. In order to test if hybrid method increases the number of peptide and protein identifications, H1299 cell lysates were analyzed with the standard, or classic method, and the hybrid method.

Both the samples were first separated with SDS-PAGE and each vertical lanes was cut into 10 bands to reduce the sample complexity. All the bands from a lane were either processed with classic method (in-gel digestion, desalting) or with the hybrid method. The data were then analyzed using X!Tandem. Importantly, the number of peptide and protein identifications were increased by 3 fold and 4.6 fold (**Figure 4.15**). So even though the separation efficiency of the hybrid method is somewhat compromised with complex mixtures, it still results in a significant increase in analytical efficiency at the mass spectrometer.

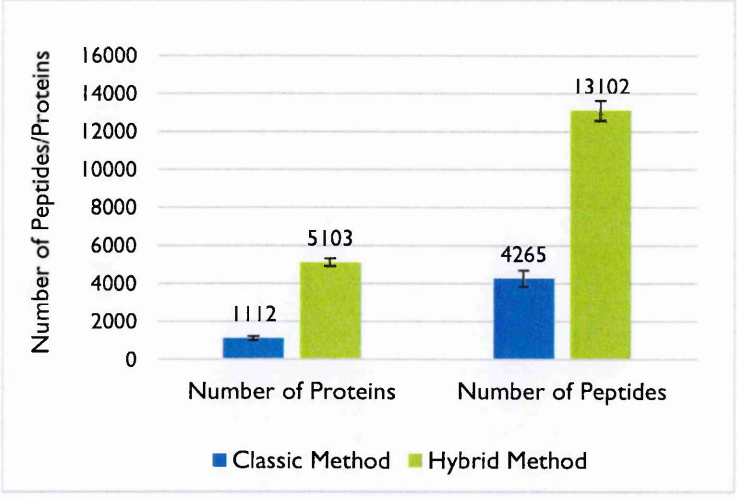


Figure 4.15. Comparison of classic method and hybrid method in proteome analysis. Cysteic acid enrichment in H1299 total cell lysates was performed as previously explained. 30 samples (10 bands and 3 fractions (flow through, was and elution)) were analyzed by LC-MS/MS and data files were merged together and analyzed as a single sample in X!Tandem. Similarly, 10 bands were analyzed for classic method and results were compared. The graphic shows the total number of identified peptides and proteins using classic method and hybrid method where cysteic acid enrichment was performed. Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.1.4.2. Decreased Peptide Overlap in Complex Samples

The significant improvements in both peptide and protein identifications are unlikely to be due to the enrichment of cysteic acids, as these peptides only account for ~14% of the identified peptides in the most enriched fraction. A key advantage of the hybrid technique

is that all the fractions are analyzed, and previously I have shown with BSA that there are very few peptides found in more than one fraction. Therefore, I also determined the overlap between fractions from complex samples. Similarly to BSA, there were very few shared peptides between fractions and the eluate contained no shared peptides **(Figure 4.16)**. This lack of shared peptides between fractions will greatly improve the analytical efficiency of the MS analysis because repeat acquisitions of the same peptide sequence will be minimized.

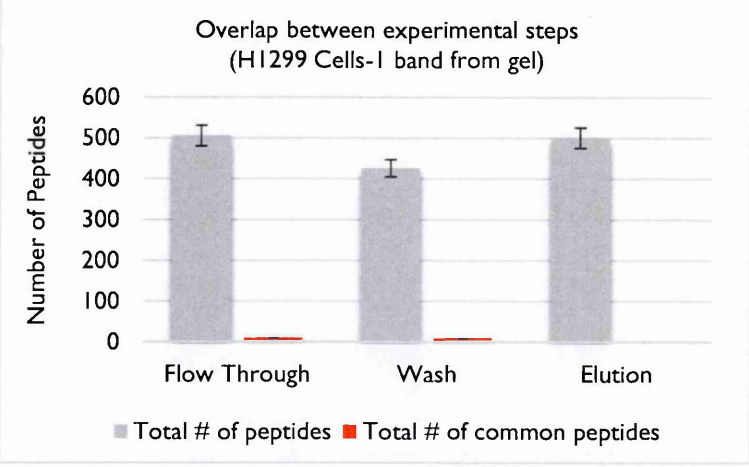


Figure 4.16. Total and common peptides that are identified in H1299 samples. H1299 cells were analyzed as previously explained. Common peptides were analyzed between three experimental steps (flow through, wash and elution). Total number of peptides and number of common peptides between experimental steps is shown on the graphic. Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.1.4.3. Increased Identification of PTMs by the Hybrid Technique

Besides increased the number of identifications, the hybrid technique also improves the identification of post translationally modified peptides. This was demonstrated by researching the original datasets to include PTMs, such as phosphorylation, oxidation, and acetylation **(Figure 4.17)**. This shows that hybrid method can also be used for studies where the primary goal is the characterization of PTMs.

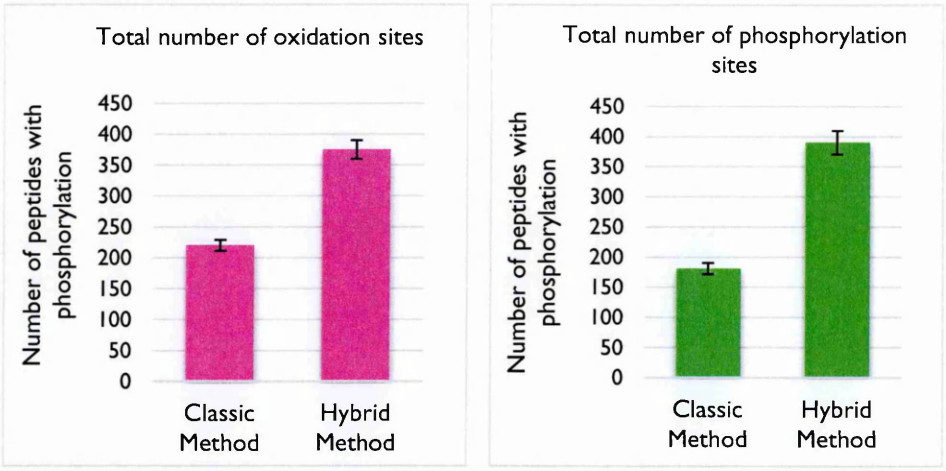


Figure 4.17. Determination of the number of identified oxidation and phosphorylation sites using the standard method or the hybrid method. H1299 cells were analyzed as previously described. For the analysis of oxidized samples, the cells were pretreated with hydrogen peroxide and the treatment with peroxide was omitted. Assays were performed in triplicate and the graphs display the mean +/- s.d.

4.1.5. Adaptation of the Hybrid Technique to Non-dividing Cells

One of the major goals of this work is the coupling of the cysteic acid purification with SILAC labeling. However, during the development of this technique, I also tried to adapt my method to systems where SILAC labeling is difficult or impossible, such as in post-mitotic cells. The initial idea was to use SILAC labeling with neuron cells, since it has been shown that non-dividing cells, such as neurons, can be tagged with SILAC amino acids in cell culture. Previously it was shown that the labeling is complete when the neurons are grown in SILAC media for at least 10 days [222]. In order to test if hybrid method can be coupled to SILAC for neuronal samples, I obtained mouse primary cortical cells and plated them in SILAC media and I kept them in culture for two weeks.

Initially I wanted to evaluate the effects of resveratrol on neurons. Interestingly, resveratrol has similar biochemical targets as sirtinol, but instead of being an HDAC inhibitor, resveratrol functions as an agonist. I started by performing viability tests on the

neuronal cultures upon resveratrol treatment. I treated primary cortical neurons separately with resveratrol or DMSO, which is the resveratrol solvent, for 24 hours. The next day I used Alamar blue to determine the viability of the cells in culture. Alamar blue oxidized by the mitochondria to produce a fluorescent compound and this conversion can be measured with a spectrophotometer or fluorimeter. The maximal concentration of DMSO used when treating cells with resveratrol was 2 mM and **Figure 4.18** indicates that neurons tolerate much higher concentrations of DMSO. Although resveratrol is typically described as a beneficial compound, a dose response curve of resveratrol in these cells showed a slight toxicity of the compound at high doses **(Figure 4.18)**.

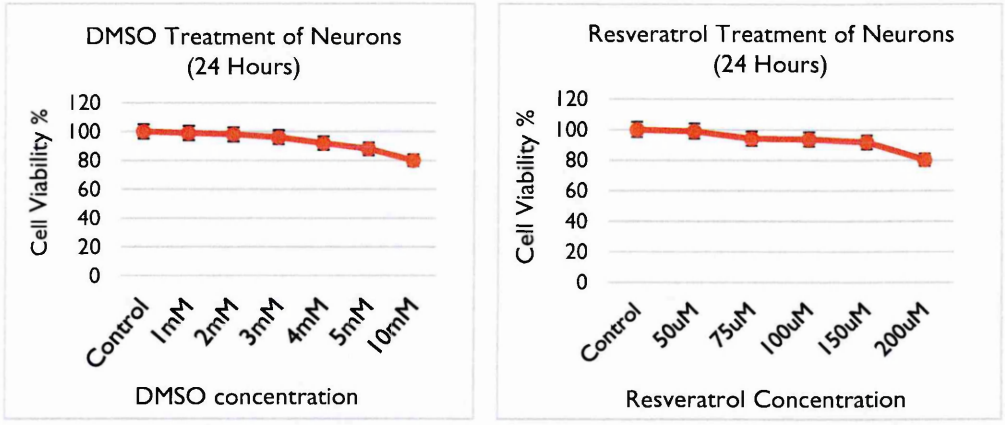
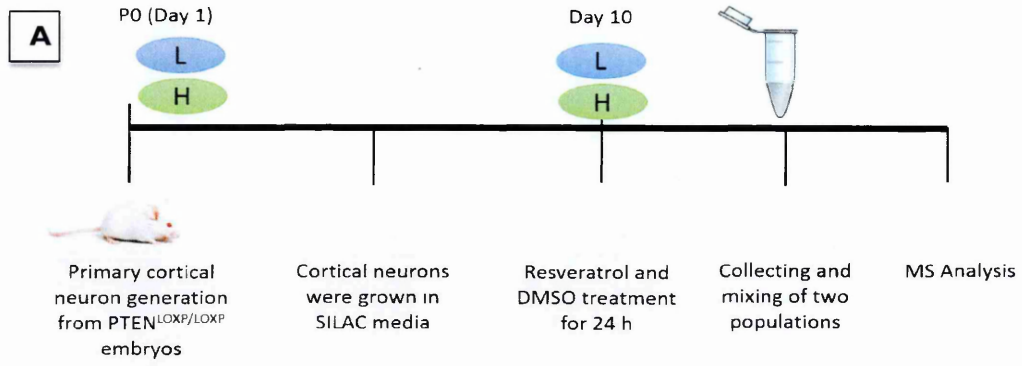


Figure 4.18. Cell viability of H1299 cells in response to resveratrol and DMSO. Cortical neurons were isolated from E13.5 mouse embryo cortex and cultured for ten days. Cells were treated with different concentrations of DMSO and resveratrol for 24 h. Cells were washed with PBS and mediums were changed. 24 h later cell viability was tested with Alamar Blue Assay. Graphic shows the percentage of cell viability with the changing concentrations of DMSO and resveratrol. Assays were performed in triplicate and the graphs display the mean viability and s.d.

Based on the viability curves, the neurons were grown in SILAC media and then treated with 150 uM resveratrol or 2 mM DMSO (control), the next day cells are collected, lysed and mixed together and processed using the hybrid method and analyzed by MS **(Figure 4.19, A)**.



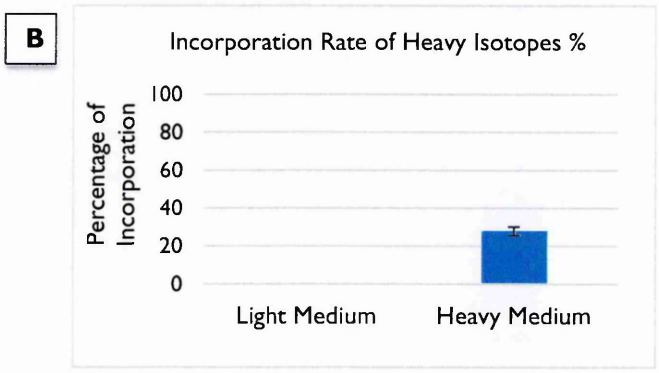


Figure 4.19. SILAC labels are poorly incorporated into cultured cortical neurons. A. Experimental scheme indicating the usage of SILAC in neuronal culturess. **B.** The rate of incorporation of the heavy isotope was determined by comparing the number of labeled lysines and arginines identified by MS. Assays were performed in triplicate and the graph displays the mean +/- s.d.

Although neurons have been previously labelled with SILAC, I have observed that the incorporation rate was quite low after 10 days (25-30 %) (**Figure 4.19, B**). When the incorporation is low, quantification rate of proteins will be unreliable, especially as it is likely that the incorporation rate for each protein depends on its own unique turnover rate. After several attempts to keep the neurons in culture for longer periods, I decided not to use SILAC for comparative analysis in neurons and instead try to adapt my sample preparation method to label free quantification methods.

Label free methods are considered less accurate than labeled methods, however, label free methods are popular because they are far easier to implement then methods that

require sample labeling. In order to test the hybrid method coupled to label-free quantitation, cultured neurons were treated with resveratrol (150 uM) or DMSO (2 mM) **(Figure 4.20, A)**. Instead of immediately mixing the samples as in the SILAC protocols, these samples were processed and analyzed separately, in this pilot experiment only a subset of the samples were analyzed by mass spectrometry. Importantly, the results showed that there is an increase in the identification rate, when samples are analyzed with hybrid method vs the classic method **(Figure 4.20, B)**. Importantly, the increased identification rate also correlated with an increase in the number of proteins that can be quantified by the hybrid technique and this improved performance is likely due to the efficient separation of peptides into distinct pools, rather than enrichment of cysteic acid.

4.2. Therapeutic Triggering of Senescence with Sirtinol in Cancer Cells

Although the initial aims of my study were to use neuronal systems, these cultures proved to be difficult and time consuming to produce. Therefore, I chose to further characterize my system using cultured tumor cell lines. Since senescence is a powerful tumor suppressor mechanism, I decided to induce senescence with sirtinol in H1299 cells, a non-small cell lung carcinoma cell line. In order to investigate the effects of sirtinol in H1299 cells, I began by characterizing their response to sirtinol. This allowed me to determine the best treatment conditions for investigating their response using the hybrid method.

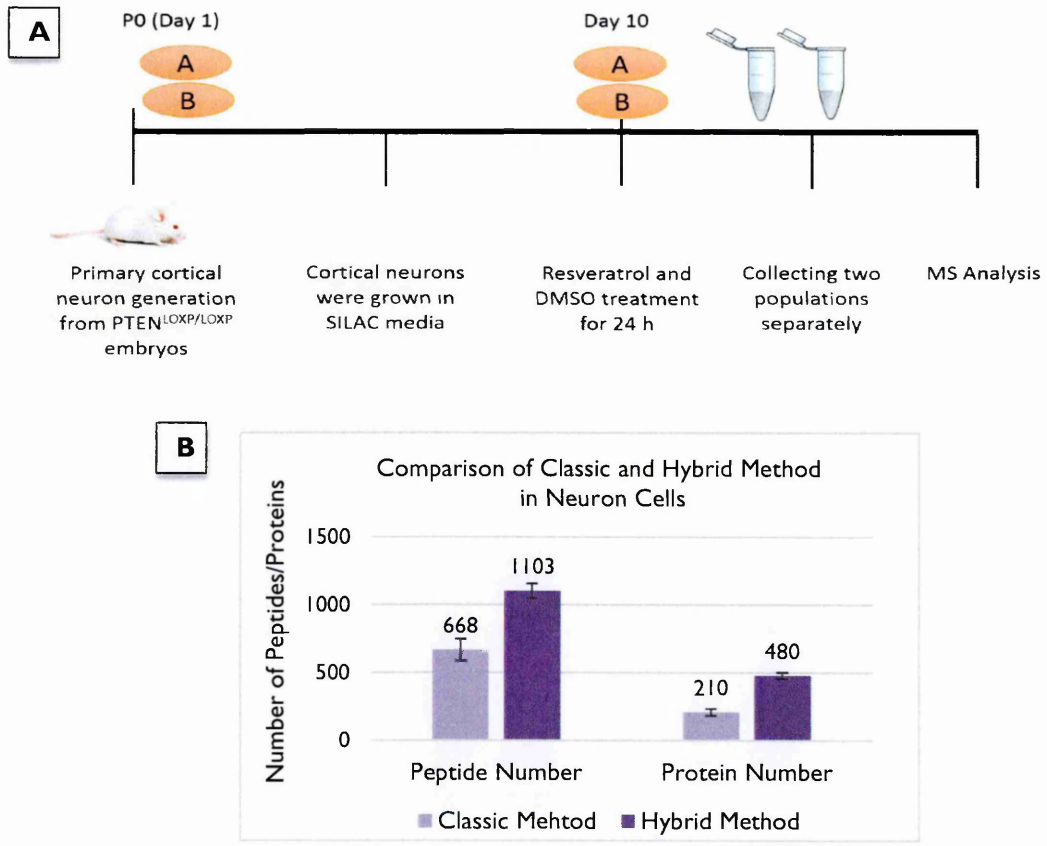


Figure 4.20. Using label-free quantification in mass spectrometry analysis of cortical neurons. **A.** Experimental scheme indicating the usage of label-free quantification in neuronal samples for mass spectrometry analysis. **B.** Comparison of peptide and protein identifications in samples analyzed by classic method and hybrid method. Only top 3 band from each samples were analyzed for comparison. Assays were performed in triplicate and the graph displays the mean +/- s.d.

4.2.1. Cell Viability of H1299 Cells Treated with Sirtinol

I performed the same viability study on H1299 cells that I performed on the cultured neurons. The H1299 cells showed a high tolerance to sirtinol treatment, because even treatment with 300 uM sirtinol resulted in only a ~20% loss in viability **(Figure 4.21)**. As there was a clear response at 300 uM sirtinol, I decided to treat H1299 cells with high concentration in the remaining experiments.

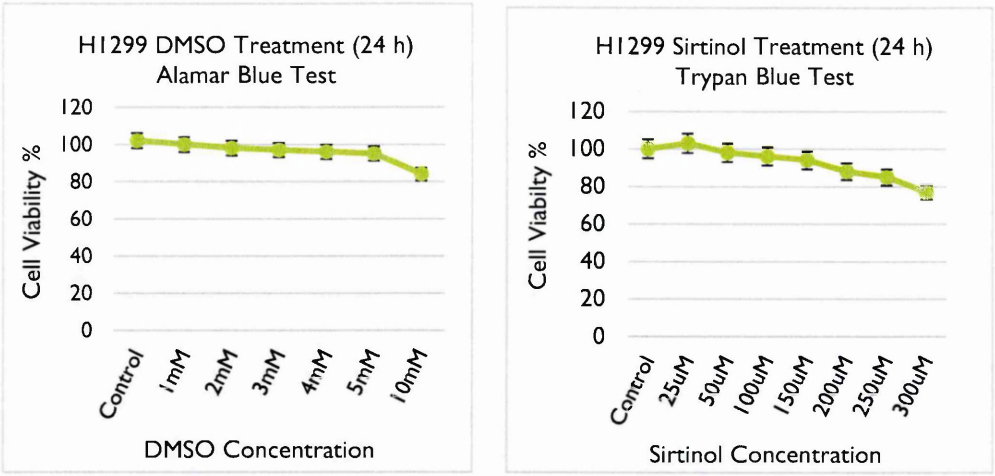


Figure 4.21. Cell viability of H1299 cells in response to sirtinol and DMSO. H1299 cells were treated with the indicate concentrations of DMSO and sirtinol for 24 h and the media changed to normal growth media. The cells were allowed to recover for another 24 h and the cell viability was determined. Assays were performed in triplicate and the graphs display the mean viability +/- s.d.

4.2.2. Sirtinol Treatment Prevents H1299 Proliferation

The previous results indicated that sirtinol is not highly toxic to H1299, and this is not too surprising as sirtinol has been shown to induce senescence rather than cell death in many tumor cell lines. Therefore, I decided to test the effects of sirtinol on H1299 proliferation. H1299 cells were treated with 200 uM or 300 uM sirtinol for 24 h and then grown for an additional 10 days. Starting from day 1 the cells were counted to check the proliferation rate. I observed that at both the concentrations of sirtinol there was a marked reduction in cell proliferation. However, 200 uM sirtinol treated cells began proliferating after 2 days in sirtinol free media. On the other hand, cell proliferation did not resume in the 300 uM sirtinol treated cells over the 10 day course of the experiment **(Figure 4.22.)**.

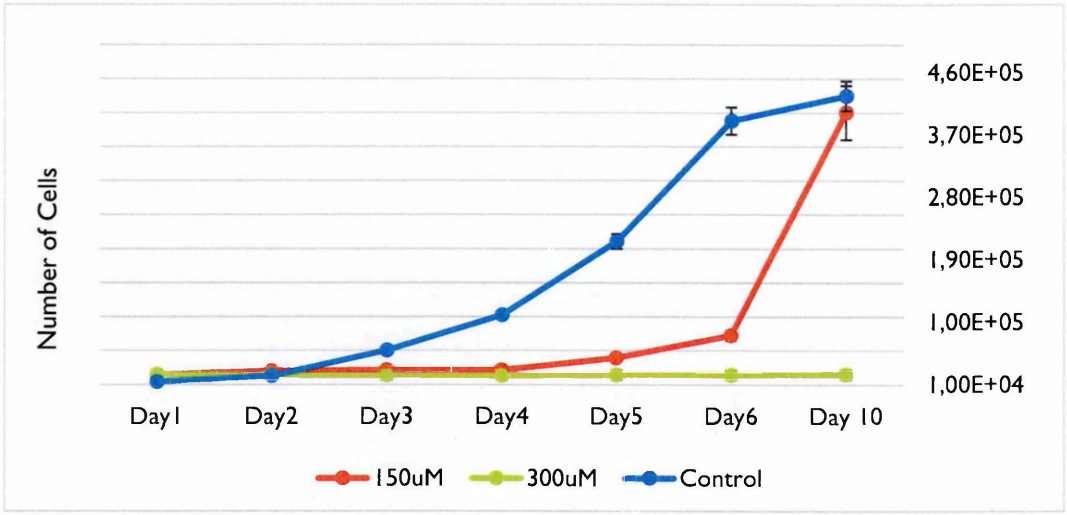


Figure 4.22. Sirtinol inhibits the proliferation of H1299 cells. H1299 cells were treated with 200 uM and 300 uM sirtinol for 24 h. Sirtinol was removed from cell and cells were cultured in inhibitor-free medium for another 10 days. Cell counts were determined and are expressed as raw cell counts +/- s.d.

4.2.3. Senescence-β-Gal Activity is Triggered in Sirtinol-Treated Cells

Cellular senescence is a permanent loss of replication even after mitogenic stimuli. In addition to the exit from the cell cycle, senescence is also characterized by the induction of a β-galactosidase activity, SA-β-gal. The SA-β-gal activity of senescent cells can be easily demonstrated by staining the cells for β-galactosidase activity using chromogenic substrates at pH 6 [108]. Since sirtinol treatment causes a prolonged inhibition of cell proliferation in H1299 cells, I decided to check the SA-β-gal activity in sirtinol treated H1299 cells, as this would demonstrate that sirtinol is inducing senescence in these cells. H1299 cells were treated with 200 uM or 300 uM sirtinol for 24 hours and then placed in sirtinol free media for 8 days. After 8 days, the presence of SA-β-gal activity was determined as described [104]. SA-β-gal activity was easily detectable in sirtinol treated cells and not surprisingly was stronger in the 300 uM treated cells.

The percentage of SA-β-Gal positive cells was approximately 90% in 300 uM sirtinol treated cells **(Figure 4.23. D).** In addition to the blue stain, many cells also became

large and flat which are also characteristic features of senescent cells **(Figure 4.23. A, B, C)**.

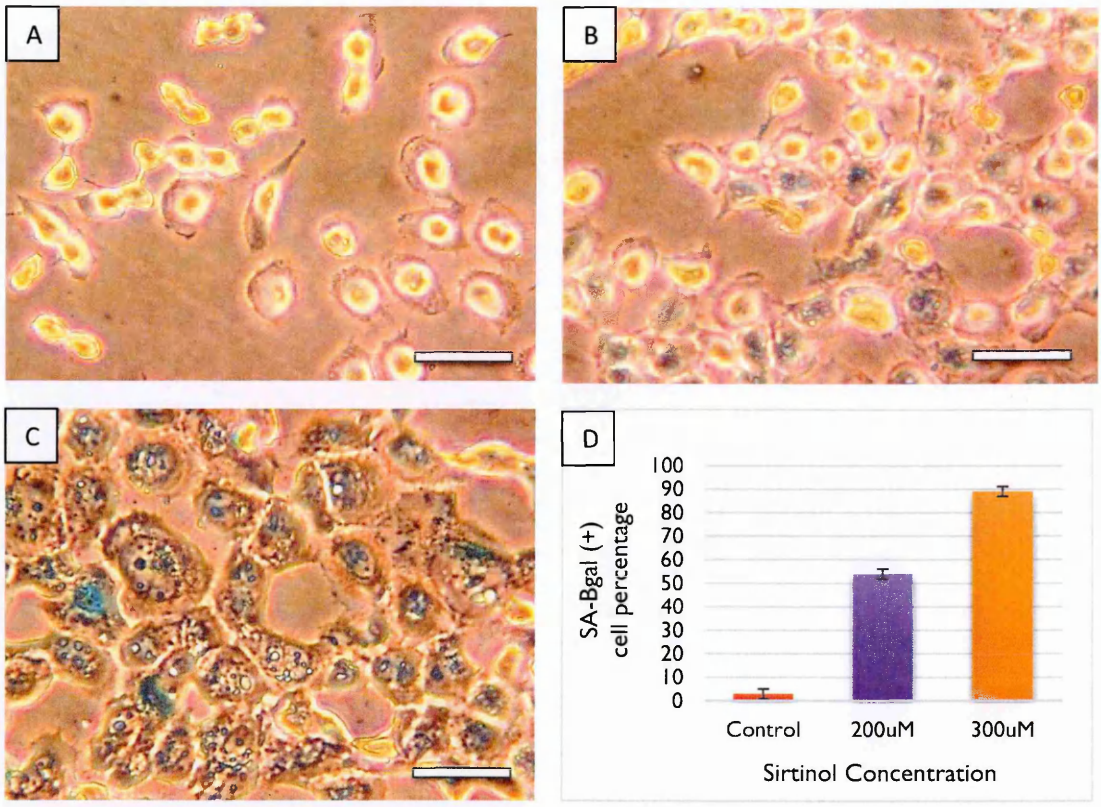


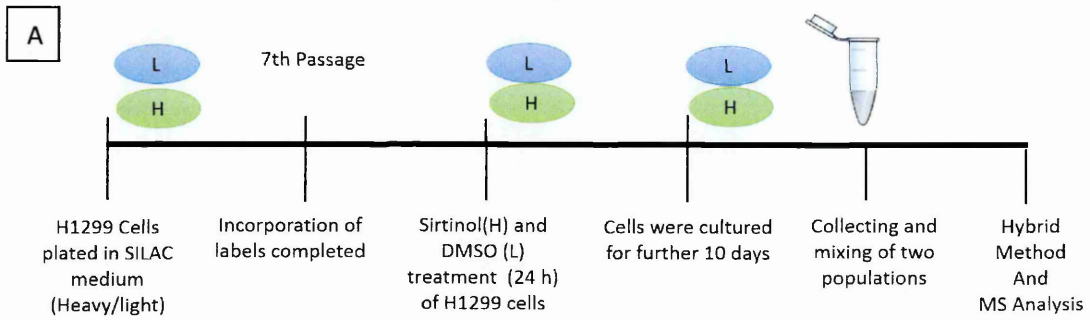
Figure 4.23. SA-β-Gal staining of sirtinol-treated H1299 cells. H1299 cells were treated with sirtinol for 24 h. The inhibitor was removed and the cells were cultured another 10 days in incubator and the SA-β-gal assay was performed. **A.** Control cells; treated with DMSO. **B.** 200 uM sirtinol treated cells. **C.** 300 uM sirtinol treated cells. **D.** Three randomly selected fields were counted and the mean percentage of SA-β-gal is shown (+/- s.d.) Scale bar = 100um.

4.3. Sirtinol Treatment Alters the Proteome of H1299 Cells

4.3.1. Proteome Analysis of Sirtinol Treated Tumor Cells

Since sirtinol causes a robust phenotype in H1299 cells, this appears to be an ideal system to further test the hybrid technique. Importantly, the phenotype is stable and does not involve cell death, which can complicate any quantitative analysis. H1299 cells were labeled with SILAC and treated with 300 uM sirtinol for 24 hours and allowed to recover for a further 8 days. The cells were lysed and mixed. The mixture was normalized to

protein amount, rather than cell number as preliminary experiments demonstrated that normalizing to cell number results in skewed isotope ratios. The mixed samples were subjected to the hybrid purification protocol and analyzed by mass spectrometry **(Figure 4.24, A)**.



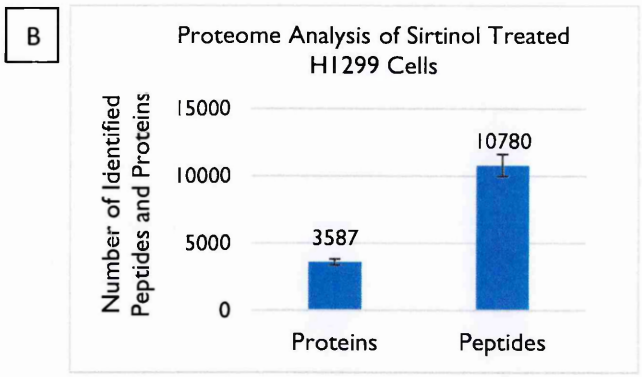


Figure 4.24. Proteome analysis of sirtinol treated H1299 tumor cells. A. Experimental scheme indicating the preparation of sirtinol treated H1299 cells for mass spectrometry analysis. **B.** Proteome analysis of sirtinol treated samples indicating the total number of peptide and protein identifications obtained from the mixed H1299 cells using the hybrid method. (+/- s.d.)

4.3.2. Sirtinol Treatment Increased Total Number of Acetylation Sites

The HDAC inhibitory activity of sirtinol is well described and is believed to be responsible for the induction of senescence in tumor cells. Therefore, sirtinol treatment is expected to result in a net increase in the number of acetylated proteins and this can be used to help determine the validity of my analytical methods. In fact, there was a marked increase in the number of identified acetylation sites in the coming from the sirtinol treated cells, with almost 200 acetylation sites identified in the Sirtinol treated samples, and 140 identified from the DMSO treated control cells **(Figure 4.25.).**

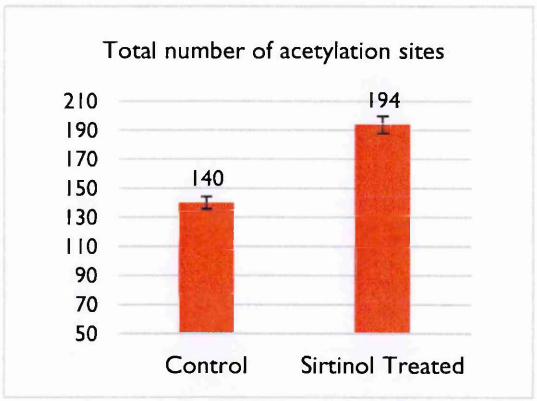


Figure 4.25. Comparison of total number of acetylation sites in control and sirtinol-treated H1299 cells. Sirtinol-treated samples were analyzed as previously explained. Graph shows the total number of acetylation sites after sirtinol treatment. (+/- s.d.)

4.3.3. Protein Expression Changes After Sirtinol treatment

Protein expression changes were obtained by analyzing the resulting data with ProteoIQ software. ProteoIQ is a comprehensive program to validate and quantify proteins and combines the results from well-known mass spectrometry database search engines such as Mascot. ProteoIQ allows you to define different groups or set of results into different grouping, including treatment groups and sample replicates. Importantly, ProteoIQ performs the quantification for SILAC labeled samples. For quantification, the RAW data

from the MS is used to produce XICs for all identified peptides. In addition, ProteoIQ performs a variety of statistical validations, such as determining the false discovery rate. ProteoIQ also uses the ProteinProphet algorithm to solve the protein inference problem.

ProteoIQ provides two methods for statistical analysis (validation) of protein and peptide identifications. Peptide or protein probability thresholds and false discovery rates are calculated to determine the confidence in peptide and protein assignments and are used to filter the data to exclude statistically insignificant identifications.

Database search results from Mascot were grouped and imported in ProteoIQ. Control and sirtinol treated H1299 samples were compared to determine the relative changes in the levels of the identified proteins **(Figure 4.26.)** The relative protein levels were determined by calculating the peak areas of the peak pairs belonging to the same peptide labeled with heavy or light isotopes and proteins with at least two unique peptide identifications were included in the analysis. In addition, a protein also needed to be identified in at least two of the repeats or it was also excluded from analysis.

As shown in **Figure 4.26**, the vast majority of the quantified proteins show no change in relative abundance following sirtinol treatment. Importantly, errors in mixing appear as most proteins exhibiting the same change in protein level and this change is the mixing error. Although, SILAC analysis allows for small changes in protein level to be detected, I decided to only focus on proteins with at least a 2 fold change in expression. This helps focus the analysis on the proteins and pathways that change the most, as well as making the downstream validation of the data easier. For example, a 20% change in protein level would be difficult to demonstrate using western blots. Using the 2 fold filter, I observed that 88 proteins were up regulated and 140 proteins were down regulated following

sirtinol treatment. The lists of up regulated and downregulated proteins are appended to the end of results section.

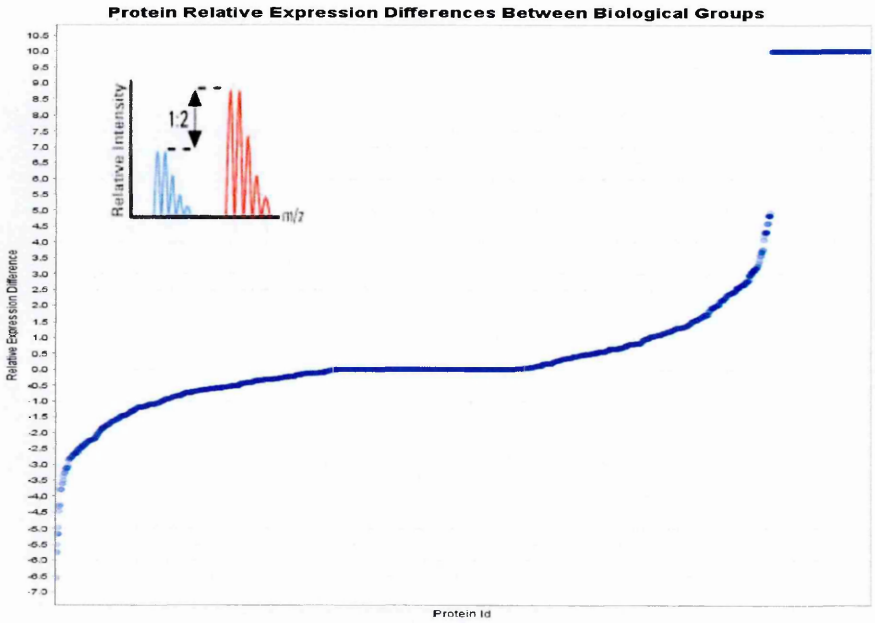


Figure 4.26. Relative protein expression levels following sirtinol treatment of H1299 cells. Relative expression levels were calculated with ProteoIQ software and sorted by magnitude of the change. The graph shows the upregulated (ascending) and downregulated (descending) proteins for all the analyzed proteins. Changes in protein levels were obtained from the measurement of different signal intensities of light (control cells) and heavy (sirtinol-treated cells) peptides (inset).

4.3.4. Validation of Proteomic Analysis with Western Blot Analysis

The ProteoIQ results revealed the proteins which were strongly upregulated or downregulated in sirtinol treated cells. I used an independent method to verify some of these changes. Since there were more than 200 proteins whose levels changed, I decided to focus on a few of these to verify the ProteoIQ analysis.

I chose two proteins which ProteoIQ indicated had reduced levels in treated samples: fatty acid synthase (FAS) and AHNAK. In the mass spectrometry results, AHNAK and FAS were found to be down regulated by 2.8 x and 3.2 x respectively. I validated the change in protein levels by western blotting for these proteins. Actin was used as a loading

control and as can be seen in **Figures 4.27. A & B,** FAS and AHNAK are downregulated by sirtinol treatment.

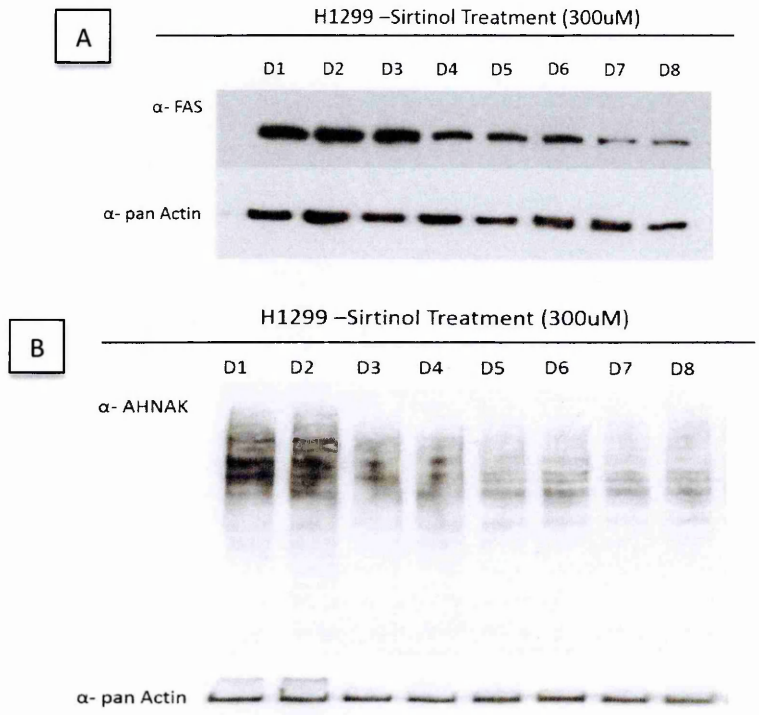


Figure 4.27. Western blot analysis of FAS and AHNAK proteins in sirtinol-treated H1299 tumor cells. A. H1299 cells were treated with 300 uM sirtinol for 24 h. Cells were collected at 1 day intervals for 8 days. Whole cell lysates were analyzed by western blotting using antibodies to FAS (A), AHNAK (B). Antibodies to pan-Actin were used as a loading control.

In addition to the downregulated proteins, I also selected two upregulated proteins for verification: eukaryotic translation initiation factor 2 alpha (EIF2α) and protein phosphatase 2 regulatory subunit B isoform R3 (PPP2R3A). These protein were upregulated by 3.2X and 3.1X respectively and western blotting was also used to validate the changes in protein levels for these proteins **(Figure 4.28. A, B.).** Although no attempt was made to verify the absolute magnitude of the change in expression, the western blot analysis reveals that ProteoIQ reliably predicts the changes in protein level following sirtinol treatment.

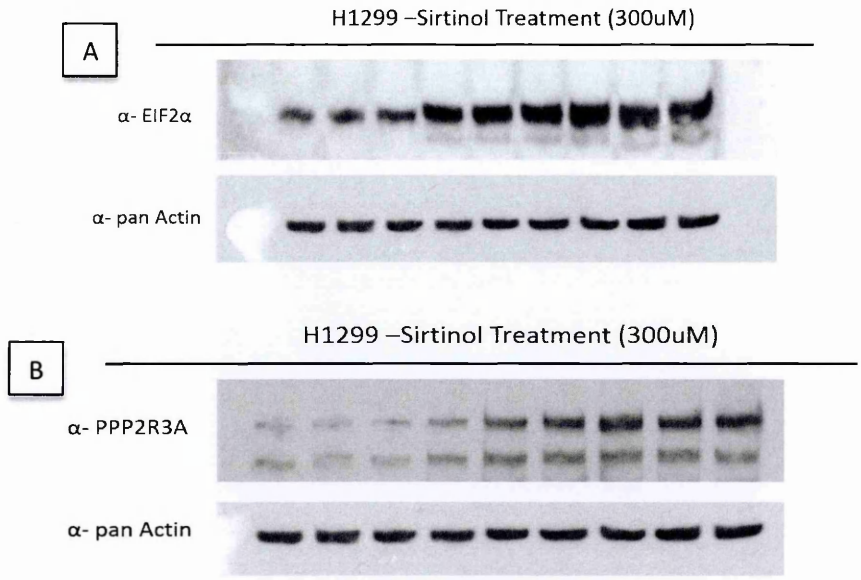


Figure 4.28. Western blot analysis of EIF2α and PPP2R3A proteins in sirtinol-treated H1299 tumor cells. A. H1299 cells were treated with 300 uM sirtinol for 24 h and the cells were collected at 1 day intervals for 8 days. Whole cell lysates were analyzed by western blotting using antibodies to EIF2 α (A) or PPP2R3A (B). Antibodies to pan-Actin were used as a loading control.

4.3.5. Investigation of Differentially Expressed Proteins with Bioinformatics Approaches

Functional annotations can be used to help determine if there are specific pathways or functions being affected by sirtinol treatment. Tools, such as DAVID (Database for Annotation, Visualization and Integrated Discovery) can be used to determine if the up or down regulated proteins consistently affect the same pathways. These typically work by extracting the gene ontology terms, or other functional annotation, for all differentially expressed proteins and then a simple statistical algorithm is used to determine significantly enriched biological functions. This type of analysis can help focus the downstream characterization of the response to those pathways that are most important.

Analysis of the ~220 proteins identified by ProteoIQ indicated that the acetylation pathway was the most strongly affected functional group following sirtinol treatment

(Figure 4.29.). Since sirtinol is known to target HDACs, this DAVID result further validates the analytical methods that I used.

Category	Term	RT	Genes	Count	%	P-Value
SP_PIR_KEYWORDS	acetylation	RT		38	65.5	6.0E-19
SP_PIR_KEYWORDS	nucleotide-binding	RT		25	43.1	1.6E-11
SP_PIR_KEYWORDS	stress response	RT		9	15.5	1.8E-11
SP_PIR_KEYWORDS	phosphoprotein	RT		42	72.4	1.7E-7
INTERPRO	Heat shock protein Hsp90, conserved site	RT		4	6.9	3.0E-7
SP_PIR_KEYWORDS	Chaperone	RT		8	13.8	4.1E-7

Figure 4.29. DAVID functional annotation chart for differentially expressed proteins in sirtinol treated H1299 cells. David software was used for determining the Gene Ontology terms that are most altered following sirtinol treatment (include reference for DAVID). As shown in the figure, the genes involved in the acetylation pathways were the most strongly affected by sirtinol treatment.

4.3.6. Long Term Effect of Sirtinol on Sirtuin Levels in H1299 Cells

Since sirtuins are involved in senescence and are considered one of the primary targets of sirtinol, I decided to also determine the effect of sirtinol treatment on sirtuin levels by western blot analysis.

Although it has been described as a direct target of sirtinol, I did not observe any significant change in Sirt1 levels (**Figure 4.30**). However, I did notice that Sirt1 is hyperphosphorylated on Ser47 in response to sirtinol (**Figure 4.30.).** Phosphorylation of Sirt1 has been suggested to regulate its activity [223] and moreover, Sirt 1 phosphorylation is also associated with senescence [224].

Interestingly, the levels of Sirt2 was found to be upregulated in response to sirtinol treatment **(Figure 4.30).** Like, with the phosphorylation of Sirt1, this may be involved in maintaining the senescent phenotype. In addition to Sirt1 and Sirt2, the protein levels of other sirtuins were found to be decreased **(Figure 4.30)**. Decreased levels of Sirt 3 [225,226], Sirt6 [227], Sirt7 [228] have also been observed in several different studies.

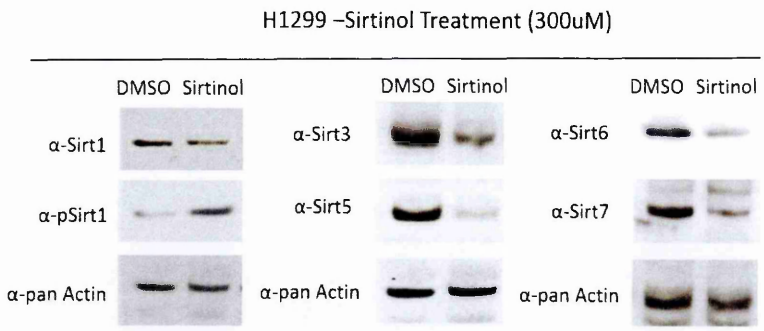


Figure 4.30. Western blot analysis of mammalian sirtuins in sirtinol treated H1299 tumor cells. H1299 cells were treated with 300 uM sirtinol for 24 h. Cells were collected after 1 and 10 days and blotted with the indicated antibodies. Antibodies to pan-Actin were used as a loading control.

4.2.7. Changes in Retinoblastoma Protein

Since sirtinol causes cell cycle arrest, I also wanted to determine if sirtinol treatment changes the level of the tumor suppressor proteins that are involved in the senescence pathways. H1299 cells have a homozygous partial deletion of the TP53 gene, hence do not express the p53 protein. Therefore, the senescent phenotype is likely to require pRb. Importantly, the activity of pRb is regulated by phosphorylation, and the hypo-phosphorylated form is the most active. In order to check if the activity level of Rb is changed, I checked Rb levels and phosphorylation status by western blot analysis. Sirtinol treatment did not change the overall levels of pRb, but, as expected, sirtinol treatment inhibits Rb phosphorylation, at three different sites (Ser780, Ser795, Ser807/811) **(Figure 4.31)**.

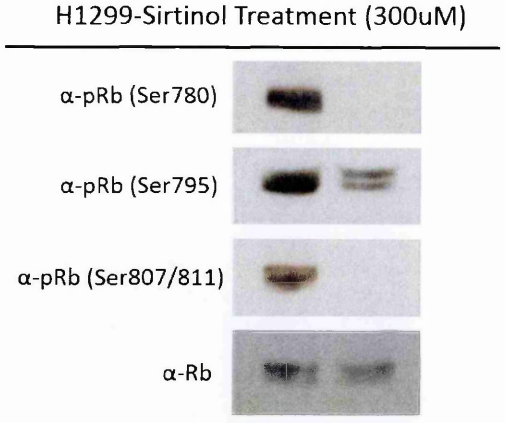


Figure 4.31. Western blot analysis of phosphorylation of retinoblastoma protein in sirtinol treated H1299 tumor cells. H1299 cells were treated with 300 uM sirtinol for 24 h. After 24 h, sirtinol was removed and cells were cultured for further 10 days. Cells were collected and lysed and blotted against pRB (Ser780), pRB (Ser795), pRB (Ser 807/811) and Rb proteins. Figure shows the changes in phosphorylation levels at three different points.

4.3.8. Comparing Results Mouse Embryonic Fibroblast

Sirtinol is a strong inducer of senescence in H1299 cells and a number of sirtinol responses were uncovered. However, it is not clear if any of these changes also occur in other forms of senescence. In order to test this, mouse embryo fibroblasts were used as they are a classic model of replicative senescence.

4.3.8.1. Obtaining Mouse Primary Fibroblast

MEFs were isolated from E13 mouse embryos. The protocol was optimized for obtaining viable MEF cultures. After a limited time in culture, 8-10 passages, MEFs undergo senescence, stop proliferating and enter a 'crisis phase' [166]. The continued serial passaging of the senescent cells eventually allows some cells to escape crisis and restart proliferative growth **(Figure 4.32. A).** These steps result in the creation of immortalized MEF cells This method is known as the 3T3 protocol and refers to passaging the cells every 3 days at a concentration of $1,5x10^4$ cells/cm^2.

In the 3T3 protocol, MEF cells start to arrest growth around 10th and 12th passages. The growth arrest is prominent by the 15th passage. This situation continues until approximately around the 23th passage. By this point, some cells have escaped crisis and they begin to take over the culture **(Figure 4.32. B)**.

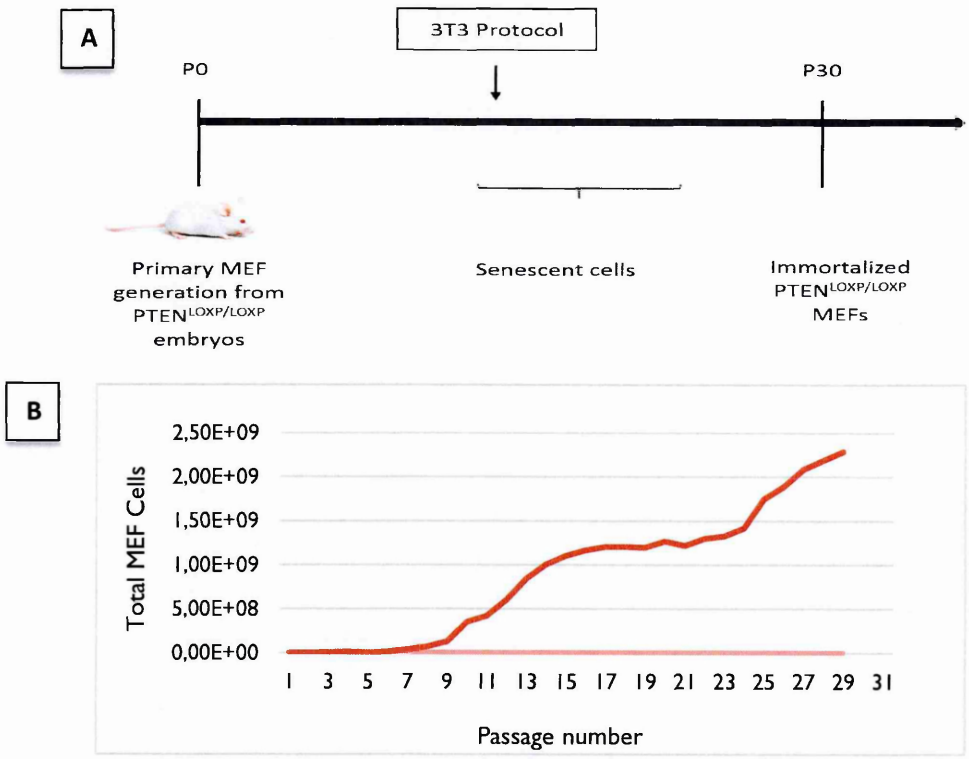


Figure 4.32. Senescence and immortalizaiton of MEF cells. A. Timeline indicating the steps necessary to make immortalized MEFs. **B.** Changes in total cell number during passages from senescence to immortalization.

4.3.8.2. Common Expression Changes between Sirtinol-Induced Senescence and Replicative Senescence

Sirtinol is a chemical inducer of senescence and may induce changes that are distinct from other senescence inducers. Therefore, I decided to test if similar changes occur during replicative senescence. MEFs were isolated and cultured under the 3T3 protocol and a portion of the cells were harvested at each passage.

Interestingly, as in sirtinol-treated tumor cells, senescent cells also down regulate in FAS. The pattern of FAS expression was clearly correlated to replicative senescence, as FAS levels started to decrease at passage 10 at the beginning of the crisis phase, and FAS levels returned as cell began exiting crisis at passage 20 **(Figure 4.33.)**.

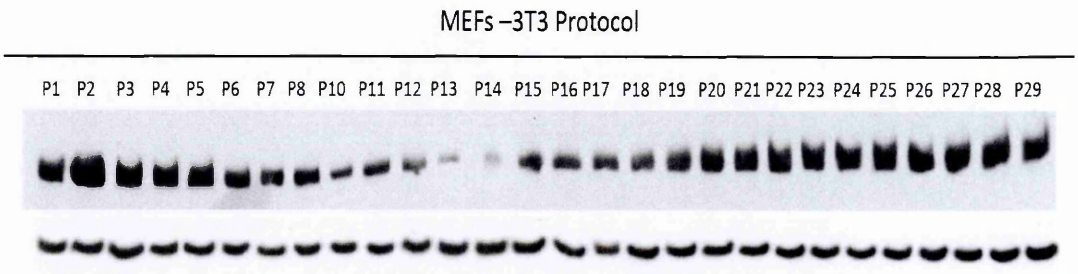


Figure 4.33. Western blot analysis of FAS protein in MEF cells. MEF cells were isolated from E13.5 mouse embryos and passaged under the 3T3 protocol. Cells were collected at each passage and blotted against FAS and pan Actin. The crisis phase is from P10 to P19.

4.3.9. Lists of Up-regulated and Down-regulated Proteins

The hybrid protocol was used to determine the changes in protein levels in response to sirtinol treatment. The changes were quantitated by the ProteoIQ software package and filtered to exclude all proteins that did not change by at least 2X. The list of down regulated proteins can be found in **Table 4.1** and the list of up-regulated proteins can be found in **Table 4.2.** Sequence ID and Sequence names are given for protein identifications. In addition the total sequence coverage, protein expression changes (ratios) and standard deviations between biological repeats are listed.

Table 4.1. List of down-regulated proteins. UniProt sequence IDs and sequence names are given for identification of proteins. Total sequence coverage (percentage of identified part of sequence in total protein sequence), protein expression ratios (C; control signal intensity, E; experimental signal intensity), C/E; expression ratio) and standard deviations (Std. Dev; measured by triplicate analysis), were given for each protein.

Sequence ID	Sequence Name	Total Sequence Coverage %	Ratio C/E	Std. Dev. %
A6NKB5	PCX2_HUMAN Pecanex-like protein 2 OS=Homo sapiens GN=PCNXL2 PE=2 SV=3	8,28	2,93	0,91
A6NM62	LRC53_HUMAN Leucine-rich repeat-containing protein 53 OS=Homo sapiens GN=LRRC53 PE=4 SV=2	17,32	3,00	1,54
A6NMY6	AXA2L_HUMAN Putative annexin A2-like protein OS=Homo sapiens GN=ANXA2P2 PE=5 SV=2	26,84	2,08	0,24
A6NNS2	DRS7C_HUMAN Dehydrogenase/reductase SDR family member 7C OS=Homo sapiens GN=DHRS7C PE=2 SV=3	8,65	2,32	1,23
A8MYB1	TMC5B_HUMAN Transmembrane and coiled-coil domain-containing protein 5B OS=Homo sapiens GN=TMCO5B PE=3 SV=1	8,79	13,76	1,51
B4DXR9	ZN732_HUMAN Zinc finger protein 732 OS=Homo sapiens GN=ZNF732 PE=2 SV=1	4,79	4,46	1,94
O00512	BCL9_HUMAN B-cell CLL/lymphoma 9 protein OS=Homo sapiens GN=BCL9 PE=1 SV=4	21,53	4,93	1,17
O43399	TPD54_HUMAN Tumor protein D54 OS=Homo sapiens GN=TPD52L2 PE=1 SV=2	30,10	4,43	2,46
O60673	DPOLZ_HUMAN DNA polymerase zeta catalytic subunit OS=Homo sapiens GN=REV3L PE=1 SV=2	14,63	4,76	0,69
O75083	WDR1_HUMAN WD repeat-containing protein 1 OS=Homo sapiens GN=WDR1 PE=1 SV=4	18,15	3,84	1,42
O75132	ZBED4_HUMAN Zinc finger BED domain-containing protein 4 OS=Homo sapiens GN=ZBED4 PE=1 SV=2	12,38	2,76	1,73
O75390	CISY_HUMAN Citrate synthase, mitochondrial OS=Homo sapiens GN=CS PE=1 SV=2	9,44	2,11	1,54
O75533	SF3B1_HUMAN Splicing factor 3B subunit 1 OS=Homo sapiens GN=SF3B1 PE=1 SV=3	12,50	1,87	0,61
O95025	SEM3D_HUMAN Semaphorin-3D OS=Homo sapiens GN=SEMA3D PE=2 SV=2	9,01	3,40	0,41

O95336	6PGL_HUMAN 6-phosphogluconolactonase OS=Homo sapiens GN=PGLS PE=1 SV=2	37,60	6,27	0,58
O95817	BAG3_HUMAN BAG family molecular chaperone regulator 3 OS=Homo sapiens GN=BAG3 PE=1 SV=3	13,39	2,04	0,80
P00505	AATM_HUMAN Aspartate aminotransferase, mitochondrial OS=Homo sapiens GN=GOT2 PE=1 SV=3	10,00	2,90	2,93
P00558	PGK1_HUMAN Phosphoglycerate kinase 1 OS=Homo sapiens GN=PGK1 PE=1 SV=3	23,98	2,35	2,16
P00568	KAD1_HUMAN Adenylate kinase isoenzyme 1 OS=Homo sapiens GN=AK1 PE=1 SV=3	16,49	2,63	2,37
P04114	APOB_HUMAN Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=2	8,83	6,64	2,28
P04843	RPN1_HUMAN Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 OS=Homo sapiens GN=RPN1 PE=1 SV=1	14,66	3,41	0,41
P06576	ATPB_HUMAN ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3	26,28	2,82	1,42
P08133	ANXA6_HUMAN Annexin A6 OS=Homo sapiens GN=ANXA6 PE=1 SV=3	16,05	5,89	1,75
P08758	ANXA5_HUMAN Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2	31,88	2,04	0,80
P09936	UCHL1_HUMAN Ubiquitin carboxyl-terminal hydrolase isozyme L1 OS=Homo sapiens GN=UCHL1 PE=1 SV=2	21,97	2,90	2,13
P11279	LAMP1_HUMAN Lysosome-associated membrane glycoprotein 1 OS=Homo sapiens GN=LAMP1 PE=1 SV=3	19,18	4,78	2,06
P12236	ADT3_HUMAN ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4	28,19	2,44	1,69
P13533	MYH6_HUMAN Myosin-6 OS=Homo sapiens GN=MYH6 PE=1 SV=5	8,46	2,52	2,86
P14678	RSMB_HUMAN Small nuclear ribonucleoprotein-associated proteins B and B' OS=Homo sapiens GN=SNRPB PE=1 SV=2	20,83	4,95	0,39
P14868	SYDC_HUMAN Aspartate--tRNA ligase, cytoplasmic OS=Homo sapiens GN=DARS PE=1 SV=2	13,17	13,66	2,07
P15822	ZEP1_HUMAN Zinc finger protein 40 OS=Homo sapiens GN=HIVEP1 PE=1 SV=3	10,01	2,15	4,89
P15924	DESP_HUMAN Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	9,93	2,69	2,47

P18085	ARF4_HUMAN ADP-ribosylation factor 4 OS=Homo sapiens GN=ARF4 PE=1 SV=3	39,44	2,45	1,13
P20929	NEBU_HUMAN Nebulin OS=Homo sapiens GN=NEB PE=1 SV=4	8,49	2,13	0,56
P22314	UBA1_HUMAN Ubiquitin-like modifier-activating enzyme 1 OS=Homo sapiens GN=UBA1 PE=1 SV=3	15,22	5,39	2,03
P23246	SFPQ_HUMAN Splicing factor, proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2	12,45	2,11	1,54
P23919	KTHY_HUMAN Thymidylate kinase OS=Homo sapiens GN=DTYMK PE=1 SV=4	10,85	2,52	4,46
P24539	AT5F1_HUMAN ATP synthase F(0) complex subunit B1, mitochondrial OS=Homo sapiens GN=ATP5F1 PE=1 SV=2	6,25	2,22	1,94
P25054	APC_HUMAN Adenomatous polyposis coli protein OS=Homo sapiens GN=APC PE=1 SV=2	15,20	6,39	0,58
P26640	SYVC_HUMAN Valine--tRNA ligase OS=Homo sapiens GN=VARS PE=1 SV=4	10,21	2,76	2,73
P27708	PYR1_HUMAN CAD protein OS=Homo sapiens GN=CAD PE=1 SV=3	4,85	2,55	1,42
P27797	CALR_HUMAN Calreticulin OS=Homo sapiens GN=CALR PE=1 SV=1	7,67	7,30	1,85
P27824	CALX_HUMAN Calnexin OS=Homo sapiens GN=CANX PE=1 SV=2	13,34	1,87	2,39
P29401	TKT_HUMAN Transketolase OS=Homo sapiens GN=TKT PE=1 SV=3	30,18	2,69	0,81
P30044	PRDX5_HUMAN Peroxiredoxin-5, mitochondrial OS=Homo sapiens GN=PRDX5 PE=1 SV=4	17,29	3,31	1,36
P30048	PRDX3_HUMAN Thioredoxin-dependent peroxide reductase, mitochondrial OS=Homo sapiens GN=PRDX3 PE=1 SV=3	22,66	6,95	1,14
P30101	PDIA3_HUMAN Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4	16,04	2,91	0,55
P35579	MYH9_HUMAN Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4	20,15	1,86	2,73
P37108	SRP14_HUMAN Signal recognition particle 14 kDa protein OS=Homo sapiens GN=SRP14 PE=1 SV=2	38,97	12,63	1,69
P38117	ETFB_HUMAN Electron transfer flavoprotein subunit beta OS=Homo sapiens GN=ETFB PE=1 SV=3	19,22	2,28	0,17

P40925	MDHC_HUMAN Malate dehydrogenase, cytoplasmic OS=Homo sapiens GN=MDH1 PE=1 SV=4	17,37	3,05	0,64
P40939	ECHA_HUMAN Trifunctional enzyme subunit alpha, mitochondrial OS=Homo sapiens GN=HADHA PE=1 SV=2	13,37	3,12	1,97
P41252	SYIC_HUMAN Isoleucine--tRNA ligase, cytoplasmic OS=Homo sapiens GN=IARS PE=1 SV=2	13,55	5,25	0,20
P42224	STAT1_HUMAN Signal transducer and activator of transcription 1-alpha/beta OS=Homo sapiens GN=STAT1 PE=1 SV=2	18,40	5,02	0,71
P42858	HD_HUMAN Huntingtin OS=Homo sapiens GN=HTT PE=1 SV=2	7,99	2,75	2,73
P49327	FAS_HUMAN Fatty acid synthase OS=Homo sapiens GN=FASN PE=1 SV=3	19,75	3,20	1,25
P49748	ACADV_HUMAN Very long-chain specific acyl-CoA dehydrogenase, mitochondrial OS=Homo sapiens GN=ACADVL PE=1 SV=1	25,04	3,65	0,66
P51148	RAB5C_HUMAN Ras-related protein Rab-5C OS=Homo sapiens GN=RAB5C PE=1 SV=2	25,93	1,97	1,31
P52272	HNRPM_HUMAN Heterogeneous nuclear ribonucleoprotein M OS=Homo sapiens GN=HNRNPM PE=1 SV=3	15,21	1,94	2,54
P54577	SYYC_HUMAN Tyrosine--tRNA ligase, cytoplasmic OS=Homo sapiens GN=YARS PE=1 SV=4	18,18	2,62	1,52
P55072	TERA_HUMAN Transitional endoplasmic reticulum ATPase OS=Homo sapiens GN=VCP PE=1 SV=4	21,84	2,39	1,31
P60174	TPIS_HUMAN Triosephosphate isomerase OS=Homo sapiens GN=TPI1 PE=1 SV=3	54,55	2,45	1,13
P60900	PSA6_HUMAN Proteasome subunit alpha type-6 OS=Homo sapiens GN=PSMA6 PE=1 SV=1	23,17	3,59	2,63
P62241	RS8_HUMAN 40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2	26,44	2,35	2,16
P62249	RS16_HUMAN 40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2	45,21	2,23	1,50
P62820	RAB1A_HUMAN Ras-related protein Rab-1A OS=Homo sapiens GN=RAB1A PE=1 SV=3	33,66	9,54	1,94
P62906	RL10A_HUMAN 60S ribosomal protein L10a OS=Homo sapiens GN=RPL10A PE=1 SV=2	13,36	5,98	1,72
P62937	PPIA_HUMAN Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2	43,64	2,08	2,62

P67809	YBOX1_HUMAN Nuclease-sensitive element-binding protein 1 OS=Homo sapiens GN=YBX1 PE=1 SV=3	22,53	2,22	1,94
P98164	LRP2_HUMAN Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	4,43	2,66	3,02
Q00577	PURA_HUMAN Transcriptional activator protein Pur-alpha OS=Homo sapiens GN=PURA PE=1 SV=2	8,07	4,69	2,27
Q00839	HNRPU_HUMAN Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6	6,79	1,86	0,81
Q02446	SP4_HUMAN Transcription factor Sp4 OS=Homo sapiens GN=SP4 PE=1 SV=2	7,40	5,22	0,20
Q09666	AHNK_HUMAN Neuroblast differentiation-associated protein AHNAK OS=Homo sapiens GN=AHNAK PE=1 SV=2	9,68	2,8	3,70
Q13634	CAD18_HUMAN Cadherin-18 OS=Homo sapiens GN=CDH18 PE=2 SV=1	10,25	1,97	0,80
Q13813	SPTN1_HUMAN Spectrin alpha chain, non-erythrocytic 1 OS=Homo sapiens GN=SPTAN1 PE=1 SV=3	8,70	2,60	1,82
Q14118	DAG1_HUMAN Dystroglycan OS=Homo sapiens GN=DAG1 PE=1 SV=2	7,04	24,27	0,44
Q14207	NPAT_HUMAN Protein NPAT OS=Homo sapiens GN=NPAT PE=1 SV=3	11,70	3,32	1,41
Q15075	EEA1_HUMAN Early endosome antigen 1 OS=Homo sapiens GN=EEA1 PE=1 SV=2	7,16	2,48	0,50
Q15365	PCBP1_HUMAN Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2	23,88	2,30	2,54
Q16555	DPYL2_HUMAN Dihydropyrimidinase-related protein 2 OS=Homo sapiens GN=DPYSL2 PE=1 SV=1	10,49	5,75	1,38
Q16650	TBR1_HUMAN T-box brain protein 1 OS=Homo sapiens GN=TBR1 PE=1 SV=1	7,77	2,01	1,21
Q16822	PCKGM_HUMAN Phosphoenolpyruvate carboxykinase [GTP], mitochondrial OS=Homo sapiens GN=PCK2 PE=1 SV=3	10,78	7,84	0,75
Q3KNS1	PTHD3_HUMAN Patched domain-containing protein 3 OS=Homo sapiens GN=PTCHD3 PE=1 SV=3	5,74	2,18	1,94
Q58FF3	ENPLL_HUMAN Putative endoplasmin-like protein OS=Homo sapiens GN=HSP90B2P PE=5 SV=1	17,79	2,43	1,69
Q5D862	FILA2_HUMAN Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1	17,06	2,81	0,62

Q5HYW2	NHSL2_HUMAN NHS-like protein 2 OS=Homo sapiens GN=NHSL2 PE=2 SV=1	27,36	2,63	3,02
Q5JPF3	AN36C_HUMAN Ankyrin repeat domain-containing protein 36C OS=Homo sapiens GN=ANKRD36C PE=2 SV=3	10,12	2,09	1,00
Q63HN8	RN213_HUMAN Isoform 2 of E3 ubiquitin-protein ligase RNF213 OS=Homo sapiens GN=RNF213	5,16	2,90	0,55
Q68DE3	K2018_HUMAN Basic helix-loop-helix domain-containing protein KIAA2018 OS=Homo sapiens GN=KIAA2018 PE=1 SV=3	6,82	5,20	0,77
Q6PCD5	RFWD3_HUMAN E3 ubiquitin-protein ligase RFWD3 OS=Homo sapiens GN=RFWD3 PE=1 SV=3	7,75	2,05	0,80
Q6PI48	SYDM_HUMAN Aspartate--tRNA ligase, mitochondrial OS=Homo sapiens GN=DARS2 PE=1 SV=1	13,02	12,30	1,95
Q6PRD1	GP179_HUMAN Probable G-protein coupled receptor 179 OS=Homo sapiens GN=GPR179 PE=1 SV=2	6,17	2,57	1,82
Q7L1T6	NB5R4_HUMAN Cytochrome b5 reductase 4 OS=Homo sapiens GN=CYB5R4 PE=1 SV=1	18,43	2,46	1,13
Q7Z494	NPHP3_HUMAN Nephrocystin-3 OS=Homo sapiens GN=NPHP3 PE=1 SV=1	7,44	3,75	4,74
Q86V15	CASZ1_HUMAN Zinc finger protein castor homolog 1 OS=Homo sapiens GN=CASZ1 PE=1 SV=4	6,42	2,29	0,17
Q86WI1	PKHL1_HUMAN Fibrocystin-L OS=Homo sapiens GN=PKHD1L1 PE=2 SV=2	7,40	4,25	1,28
Q8IVF2	AHNK2_HUMAN Protein AHNAK2 OS=Homo sapiens GN=AHNAK2 PE=1 SV=2	7,90	1,99	1,20
Q8IWJ2	GCC2_HUMAN GRIP and coiled-coil domain-containing protein 2 OS=Homo sapiens GN=GCC2 PE=1 SV=4	12,35	5,17	0,71
Q8IZL2	MAML2_HUMAN Mastermind-like protein 2 OS=Homo sapiens GN=MAML2 PE=1 SV=2	22,49	2,71	0,02
Q8N0V5	GNT2A_HUMAN N-acetyllactosaminide beta-1,6-N-acetylglucosaminyl-transferase, isoform A OS=Homo sapiens GN=GCNT2 PE=2 SV=1	15,67	3,93	0,07
Q8N442	GUF1_HUMAN Translation factor GUF1, mitochondrial OS=Homo sapiens GN=GUF1 PE=1 SV=1	8,07	25,66	2,40
Q8NCM8	DYHC2_HUMAN Cytoplasmic dynein 2 heavy chain 1 OS=Homo sapiens GN=DYNC2H1 PE=1 SV=4	9,52	2,32	1,05
Q8NDH2	CC168_HUMAN Coiled-coil domain-containing protein 168 OS=Homo sapiens GN=CCDC168 PE=2 SV=2	10,73	2,93	0,91

Q8NF91	SYNE1_HUMAN Nesprin-1 OS=Homo sapiens GN=SYNE1 PE=1 SV=4	6,81	3,23	0,68
Q8WUM4	PDC6I_HUMAN Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1	5,88	1,89	1,08
Q92889	XPF_HUMAN DNA repair endonuclease XPF OS=Homo sapiens GN=ERCC4 PE=1 SV=3	5,24	3,56	1,23
Q92900	RENT1_HUMAN Regulator of nonsense transcripts 1 OS=Homo sapiens GN=UPF1 PE=1 SV=2	8,24	3,05	0,64
Q96C10	DHX58_HUMAN Probable ATP-dependent RNA helicase DHX58 OS=Homo sapiens GN=DHX58 PE=1 SV=1	12,98	2,26	1,50
Q96SI9	STRBP_HUMAN Spermatid perinuclear RNA-binding protein OS=Homo sapiens GN=STRBP PE=1 SV=1	5,36	5,26	2,03
Q96T58	MINT_HUMAN Msx2-interacting protein OS=Homo sapiens GN=SPEN PE=1 SV=1	5,21	5,73	1,38
Q96TA1	NIBL1_HUMAN Niban-like protein 1 OS=Homo sapiens GN=FAM129B PE=1 SV=3	3,89	2,80	0,62
Q9BYK8	HELZ2_HUMAN Helicase with zinc finger domain 2 OS=Homo sapiens GN=HELZ2 PE=1 SV=6	4,27	3,26	0,68
Q9BZZ5	API5_HUMAN Apoptosis inhibitor 5 OS=Homo sapiens GN=API5 PE=1 SV=3	12,98	3,10	1,97
Q9C0J8	WDR33_HUMAN pre-mRNA 3' end processing protein WDR33 OS=Homo sapiens GN=WDR33 PE=1 SV=2	16,99	4,77	4,74
Q9H0K1	SIK2_HUMAN Serine/threonine-protein kinase SIK2 OS=Homo sapiens GN=SIK2 PE=1 SV=1	4,97	3,81	1,17
Q9H2Y7	ZN106_HUMAN Zinc finger protein 106 OS=Homo sapiens GN=ZNF106 PE=1 SV=1	7,97	2,43	0,60
Q9H4A6	GOLP3_HUMAN Golgi phosphoprotein 3 OS=Homo sapiens GN=GOLPH3 PE=1 SV=1	7,38	4,48	1,94
Q9H6S0	YTDC2_HUMAN Probable ATP-dependent RNA helicase YTHDC2 OS=Homo sapiens GN=YTHDC2 PE=1 SV=2	13,08	3,38	0,41
Q9HDC5	JPH1_HUMAN Junctophilin-1 OS=Homo sapiens GN=JPH1 PE=1 SV=2	7,56	2,36	1,88
Q9NS40	KCNH7_HUMAN Potassium voltage-gated channel subfamily H member 7 OS=Homo sapiens GN=KCNH7 PE=2 SV=2	11,45	5,43	0,26
Q9NVH0	EXD2_HUMAN Exonuclease 3'-5' domain-containing protein 2 OS=Homo sapiens GN=EXD2 PE=1 SV=2	6,12	12,02	1,95

Q9NXZ1	SAGE1_HUMAN Sarcoma antigen 1 OS=Homo sapiens GN=SAGE1 PE=1 SV=2	12,17	5,19	0,77
Q9NYQ6	CELR1_HUMAN Cadherin EGF LAG seven-pass G-type receptor 1 OS=Homo sapiens GN=CELSR1 PE=1 SV=1	9,19	1,97	0,09
Q9NYW7	TA2R1_HUMAN Taste receptor type 2 member 1 OS=Homo sapiens GN=TAS2R1 PE=1 SV=1	3,68	2,81	0,62
Q9P031	TAP26_HUMAN Thyroid transcription factor 1-associated protein 26 OS=Homo sapiens GN=CCDC59 PE=1 SV=2	11,62	3,21	1,33
Q9P2E3	ZNFX1_HUMAN NFX1-type zinc finger-containing protein 1 OS=Homo sapiens GN=ZNFX1 PE=1 SV=2	3,44	2,87	0,05
Q9P2J5	SYLC_HUMAN Leucine--tRNA ligase, cytoplasmic OS=Homo sapiens GN=LARS PE=1 SV=2	10,37	1,97	0,80
Q9UPN3	MACF1_HUMAN Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5 OS=Homo sapiens GN=MACF1 PE=1 SV=4	5,51	1,87	1,08
Q9UQ35	SRRM2_HUMAN Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	8,94	6,19	1,58
Q9UQ80	PA2G4_HUMAN Proliferation-associated protein 2G4 OS=Homo sapiens GN=PA2G4 PE=1 SV=3	24,62	2,63	3,02
Q9Y2W1	TR150_HUMAN Thyroid hormone receptor-associated protein 3 OS=Homo sapiens GN=THRAP3 PE=1 SV=2	5,45	4,66	2,27
Q9Y5S2	MRCKB_HUMAN Serine/threonine-protein kinase MRCK beta OS=Homo sapiens GN=CDC42BPB PE=1 SV=2	9,12	1,96	0,04
Q9Y696	CLIC4_HUMAN Chloride intracellular channel protein 4 OS=Homo sapiens GN=CLIC4 PE=1 SV=4	24,51	2,26	1,50
Q9Y6L7	TLL2_HUMAN Tolloid-like protein 2 OS=Homo sapiens GN=TLL2 PE=1 SV=1	4,83	3,66	1,10
Q9Y6N5	SQRD_HUMAN Sulfide:quinone oxidoreductase, mitochondrial OS=Homo sapiens GN=SQRDL PE=1 SV=1	16,22	3,53	2,04

Table 4.2. List of up-regulated proteins. UniProt sequence IDs and sequence names are given for identification of proteins. Total sequence coverage (percentage of identified part of sequence in total protein sequence), protein expression ratios (C; control signal intensity, E; experimental signal intensity), E/C; expression ratio) and standard deviations (Std. Dev; measured by triplicate analysis), were given for each protein.

Sequence ID	Sequence Name	Total Sequence Coverage %	Ratio	Std. Dev. %
O00571	DDX3X_HUMAN ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3	10,25	2,39	0,89
O15067	PUR4_HUMAN Phosphoribosylformylglycinamidine synthase OS=Homo sapiens GN=PFAS PE=1 SV=4	6,65	3,73	0,99
O43175	SERA_HUMAN D-3-phosphoglycerate dehydrogenase OS=Homo sapiens GN=PHGDH PE=1 SV=4	8,13	2,90	1,02
O43181	NDUS4_HUMAN NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial OS=Homo sapiens GN=NDUFS4 PE=1 SV=1	9,64	3,10	1,06
O60503	ADCY9_HUMAN Adenylate cyclase type 9 OS=Homo sapiens GN=ADCY9 PE=1 SV=4	9,5	3,33	1,02
O60741	HCN1_HUMAN Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 1 OS=Homo sapiens GN=HCN1 PE=2 SV=3	6,2	3,41	0,99
O60811	PRAM2_HUMAN PRAME family member 2 OS=Homo sapiens GN=PRAMEF2 PE=2 SV=2	7,3	2,06	0,79
O75486	SUPT3_HUMAN Transcription initiation protein SPT3 homolog OS=Homo sapiens GN=SUPT3H PE=1 SV=2	9,1	2,53	0,65
O75531	BAF_HUMAN Barrier-to-autointegration factor OS=Homo sapiens GN=BANF1 PE=1 SV=1	10,2	5,67	1,12
O76094	SRP72_HUMAN Signal recognition particle subunit SRP72 OS=Homo sapiens GN=SRP72 PE=1 SV=3	8,25	1,95	1,09
O94927	HAUS5_HUMAN HAUS augmin-like complex subunit 5 OS=Homo sapiens GN=HAUS5 PE=1 SV=2	6,14	2,96	1,2
O95425	SVIL_HUMAN Supervillin OS=Homo sapiens GN=SVIL PE=1 SV=2	6,07	5,25	0,95
P06753	TPM3_HUMAN Tropomyosin alpha-3 chain OS=Homo sapiens GN=TPM3 PE=1 SV=2	10,53	3,01	0,63

P07900	HS90A_HUMAN Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5	11,36	2,10	0,8
P10412	H14_HUMAN Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	38,81	4,69	1,44
P16104	H2AX_HUMAN Histone H2AX OS=Homo sapiens GN=H2AFX PE=1 SV=2	52,45	2,01	2,63
P16401	H15_HUMAN Histone H1.5 OS=Homo sapiens GN=HIST1H1B PE=1 SV=3	28,32	2,24	0,79
P16402	H13_HUMAN Histone H1.3 OS=Homo sapiens GN=HIST1H1D PE=1 SV=2	24,43	4,14	1,45
P16403	H12_HUMAN Histone H1.2 OS=Homo sapiens GN=HIST1H1C PE=1 SV=2	22,54	4,14	1,45
P18621	RL17_HUMAN 60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3	38,04	2,58	3,49
P20042	IF2B_HUMAN Eukaryotic translation initiation factor 2 subunit 2 OS=Homo sapiens GN=EIF2S2 PE=1 SV=2	10,81	3,1	2,62
P21333	FLNA_HUMAN Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4	19,25	1,90	1,25
P23396	RS3_HUMAN 40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2	32,51	2,13	0,46
P30414	NKTR_HUMAN NK-tumor recognition protein OS=Homo sapiens GN=NKTR PE=1 SV=2	9,87	3,06	1,26
P31629	ZEP2_HUMAN Transcription factor HIVEP2 OS=Homo sapiens GN=HIVEP2 PE=1 SV=2	10,36	3,11	1,16
P31939	PUR9_HUMAN Bifunctional purine biosynthesis protein PURH OS=Homo sapiens GN=ATIC PE=1 SV=3	10,20	2,51	1,65
P35555	FBN1_HUMAN Fibrillin-1 OS=Homo sapiens GN=FBN1 PE=1 SV=3	28,10	11,70	1,87
P40227	TCPZ_HUMAN T-complex protein 1 subunit zeta OS=Homo sapiens GN=CCT6A PE=1 SV=3	15,60	2,10	0,99
P42345	MTOR_HUMAN Serine/threonine-protein kinase mTOR OS=Homo sapiens GN=MTOR PE=1 SV=1	9,30	2,02	0,63
P49368	TCPG_HUMAN T-complex protein 1 subunit gamma OS=Homo sapiens GN=CCT3 PE=1 SV=4	15,90	1,93	1,12
P49593	PPM1F_HUMAN Protein phosphatase 1F OS=Homo sapiens GN=PPM1F PE=1 SV=3	12,30	8,14	0,95

P55884	EIF3B_HUMAN Eukaryotic translation initiation factor 3 subunit B OS=Homo sapiens GN=EIF3B PE=1 SV=3	14,10	2,15	0,63
P56715	RP1_HUMAN Oxygen-regulated protein 1 OS=Homo sapiens GN=RP1 PE=1 SV=1	9,25	5,67	0,97
P60842	IF4A1_HUMAN Eukaryotic initiation factor 4A-I OS=Homo sapiens GN=EIF4A1 PE=1 SV=1	20,44	1,92	0,30
P61247	RS3A_HUMAN 40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2	14,39	2,83	0,86
P62805	H4_HUMAN Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2	45,63	2,13	1,62
P62829	RL23_HUMAN 60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1 SV=1	15,23	2,16	1,20
P62917	RL8_HUMAN 60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2	14,79	1,79	1,39
Q01813	K6PP_HUMAN 6-phosphofructokinase type C OS=Homo sapiens GN=PFKP PE=1 SV=2	9,45	2,18	1,02
Q01959	SC6A3_HUMAN Sodium-dependent dopamine transporter OS=Homo sapiens GN=SLC6A3 PE=1 SV=1	8,17	3,11	1,30
Q06190	P2R3A_HUMAN Serine/threonine-protein phosphatase 2A regulatory subunit B" subunit alpha OS=Homo sapiens GN=PPP2R3A PE=1 SV=1	8,11	3,2	2,11
Q07954	LRP1_HUMAN Prolow-density lipoprotein receptor-related protein 1 OS=Homo sapiens GN=LRP1 PE=1 SV=2	8,20	3,68	0,78
Q12888	TP53B_HUMAN Tumor suppressor p53-binding protein 1 OS=Homo sapiens GN=TP53BP1 PE=1 SV=2	16,63	2,32	1,23
Q13201	MMRN1_HUMAN Multimerin-1 OS=Homo sapiens GN=MMRN1 PE=1 SV=3 not in gPM	8,63	3,67	0,95
Q13283	G3BP1_HUMAN Ras GTPase-activating protein-binding protein 1 OS=Homo sapiens GN=G3BP1 PE=1 SV=1	8,57	6,29	1,06
Q14152	EIF3A_HUMAN Eukaryotic translation initiation factor 3 subunit A OS=Homo sapiens GN=EIF3A PE=1 SV=1	5,35	1,94	1,09
Q14240	IF4A2_HUMAN Eukaryotic initiation factor 4A-II OS=Homo sapiens GN=EIF4A2 PE=1 SV=2	21,62	2,02	0,17
Q14258	TRI25_HUMAN E3 ubiquitin/ISG15 ligase TRIM25 OS=Homo sapiens GN=TRIM25 PE=1 SV=2	9,98	1,94	1,11
Q14315	FLNC_HUMAN Filamin-C OS=Homo sapiens GN=FLNC PE=1 SV=3	30,17	1,93	1,21

Q14573	ITPR3_HUMAN Inositol 1,4,5-trisphosphate receptor type 3 OS=Homo sapiens GN=ITPR3 PE=1 SV=2	15,60	3,93	1,30
Q14847	LASP1_HUMAN LIM and SH3 domain protein 1 OS=Homo sapiens GN=LASP1 PE=1 SV=2	16,86	1,94	2,35
Q15596	NCOA2_HUMAN Nuclear receptor coactivator 2 OS=Homo sapiens GN=NCOA2 PE=1 SV=2	13,69	2,00	1,21
Q15751	HERC1_HUMAN Probable E3 ubiquitin-protein ligase HERC1 OS=Homo sapiens GN=HERC1 PE=1 SV=2 3 peptides loge -11	11,25	2,90	0,57
Q16099	GRIK4_HUMAN Glutamate receptor ionotropic, kainate 4 OS=Homo sapiens GN=GRIK4 PE=2 SV=2	10,40	2,19	0,95
Q16222	UAP1_HUMAN UDP-N-acetylhexosamine pyrophosphorylase OS=Homo sapiens GN=UAP1 PE=1 SV=3	9,96	3,52	0,98
Q16880	CGT_HUMAN 2-hydroxyacylsphingosine 1-beta-galactosyltransferase OS=Homo sapiens GN=UGT8 PE=2 SV=2	8,95	3,73	1,14
Q4G0N4	NAKD2_HUMAN NAD kinase 2, mitochondrial OS=Homo sapiens GN=NADK2 PE=1 SV=2	14,71	2,47	0,57
Q52LR7	EPC2_HUMAN Enhancer of polycomb homolog 2 OS=Homo sapiens GN=EPC2 PE=1 SV=2	12,78	2,65	0,95
Q68EM7	RHG17_HUMAN Rho GTPase-activating protein 17 OS=Homo sapiens GN=ARHGAP17 PE=1 SV=1	10,63	3,19	1,14
Q7Z4L5	TT21B_HUMAN Tetratricopeptide repeat protein 21B OS=Homo sapiens GN=TTC21B PE=1 SV=2	9,98	2,02	1,05
Q8IUE6	H2A2B_HUMAN Histone H2A type 2-B OS=Homo sapiens GN=HIST2H2AB PE=1 SV=3	52,31	1,90	1,09
Q8IZT6	ASPM_HUMAN Abnormal spindle-like microcephaly-associated protein OS=Homo sapiens GN=ASPM PE=1 SV=2	6,79	7,13	1,12
Q8N394	TMTC2_HUMAN Transmembrane and TPR repeat-containing protein 2 OS=Homo sapiens GN=TMTC2 PE=2 SV=1	10,20	3,05	1,14
Q8N4B4	FBX39_HUMAN F-box only protein 39 OS=Homo sapiens GN=FBXO39 PE=2 SV=1	6,33	5,30	0,99
Q8N7X1	RMXL3_HUMAN RNA-binding motif protein, X-linked-like-3 OS=Homo sapiens GN=RBMXL3 PE=2 SV=2	10,86	2,35	0,57
Q8TBJ5	FEZF2_HUMAN Fez family zinc finger protein 2 OS=Homo sapiens GN=FEZF2 PE=2 SV=2	10,23	4,37	1,25
Q92574	TSC1_HUMAN Hamartin OS=Homo sapiens GN=TSC1 PE=1 SV=2	8,67	3,53	1,12

Q969M2	CXA10_HUMAN Gap junction alpha-10 protein OS=Homo sapiens GN=GJA10 PE=2 SV=1	9,68	3,34	1,04
Q96K78	GP128_HUMAN Probable G-protein coupled receptor 128 OS=Homo sapiens GN=GPR128 PE=2 SV=2	8,71	4,66	1,09
Q96M42	CU129_HUMAN Putative uncharacterized protein encoded by LINC00479 OS=Homo sapiens GN=LINC00479 PE=5 SV=2	28,87	3,63	1,12
Q96QV6	H2A1A_HUMAN Histone H2A type 1-A OS=Homo sapiens GN=HIST1H2AA PE=1 SV=3	35,11	3,44	1,12
Q99541	PLIN2_HUMAN Perilipin-2 OS=Homo sapiens GN=PLIN2 PE=1 SV=2	25,6	5,58	0,57
Q99832	TCPH_HUMAN T-complex protein 1 subunit eta OS=Homo sapiens GN=CCT7 PE=1 SV=2	21,36	2,03	0,95
Q99880	H2B1L_HUMAN Histone H2B type 1-L OS=Homo sapiens GN=HIST1H2BL PE=1 SV=3	34,13	3,14	0,23
Q9BYE9	CDHR2_HUMAN Cadherin-related family member 2 OS=Homo sapiens GN=CDHR2 PE=1 SV=2	23,65	2,15	0,86
Q9BYF1	ACE2_HUMAN Angiotensin-converting enzyme 2 OS=Homo sapiens GN=ACE2 PE=1 SV=2	23,73	3,55	2,01
Q9H2T7	RBP17_HUMAN Ran-binding protein 17 OS=Homo sapiens GN=RANBP17 PE=2 SV=1	9,56	2,71	1,01
Q9H583	HEAT1_HUMAN HEAT repeat-containing protein 1 OS=Homo sapiens GN=HEATR1 PE=1 SV=3	15,69	5,67	1,04
Q9NXV6	CARF_HUMAN CDKN2A-interacting protein OS=Homo sapiens GN=CDKN2AIP PE=1 SV=3	20,30	2,13	1,15
Q9NZI8	IF2B1_HUMAN Insulin-like growth factor 2 mRNA-binding protein 1 OS=Homo sapiens GN=IGF2BP1 PE=1 SV=2	13,17	2,18	1,26
Q9P2P5	HECW2_HUMAN E3 ubiquitin-protein ligase HECW2 OS=Homo sapiens GN=HECW2 PE=1 SV=2	18,4	3,05	1,13
Q9UKN1	MUC12_HUMAN Mucin-12 OS=Homo sapiens GN=MUC12 PE=1 SV=2	12,5	6,56	1,26
Q9UL63	MKLN1_HUMAN Muskelin OS=Homo sapiens GN=MKLN1 PE=1 SV=2	14,63	10,30	1,21
Q9UM73	ALK_HUMAN ALK tyrosine kinase receptor OS=Homo sapiens GN=ALK PE=1 SV=3	15,69	16,00	1,14
Q9Y262	EIF3L_HUMAN Eukaryotic translation initiation factor 3 subunit L OS=Homo sapiens GN=EIF3L PE=1 SV=1	8,95	3,38	1,09

Q9Y3S1	WNK2_HUMAN Serine/threonine-protein kinase WNK2 OS=Homo sapiens GN=WNK2 PE=1 SV=4	4,53	1,91	2,46
Q9Y5Y6	ST14_HUMAN Suppressor of tumorigenicity 14 protein OS=Homo sapiens GN=ST14 PE=1 SV=2	16,3	2,12	1,02
Q9Y617	SERC_HUMAN Phosphoserine aminotransferase OS=Homo sapiens GN=PSAT1 PE=1 SV=2	20,54	2,10	1,13

5. CONCLUSIONS

Proteomics is a rapidly growing field of biology that is used for the study of proteins. Since proteins regulate and mediate virtually all cellular functions, proteomic studies have been used for diverse research areas including drug discovery and systems biology. Information gained from proteomics studies is particularly useful in helping to understand cellular mechanisms, especially because the proteome can be altered in many ways that cannot be predicted from genetic analysis.

Typically proteomics studies attempt to identify or characterize all the proteins that are found in a particular biological context. For example, using proteomics to identify the changes in protein levels in response to different stimulants will lead to a better understanding of how the stimulant functions. However, analyzing the alterations in a proteome is highly dependent on choosing the right techniques and applying them in a way that maximizes their analytical power.

Although a large number of proteomics techniques have been developed, there are still many areas where even the most advanced analysis are insufficient. In this regard, further improvements are needed to increase the sensitivity and throughput of proteomics analysis. Unlike genomic analysis, it is currently impossible to identify all the proteins found in a complex biological fluid, such as serum. With this perspective, one of the major goals of my thesis was to develop a method for mass spectrometry analysis that improves the number of peptide and protein identifications from a sample. Importantly, I also thought a method that was easily adaptable to quantitative studies.

In the second part, I used the method I developed, referred to here as the "hybrid method" to investigate the changes in the proteome of cancer cells to sirtinol. The second part of my thesis gives information on the relationship between cancer and senescence.

5.1. Advantages of Hybrid Method

The method I developed, which I refer to as the hybrid method, uses classical SILAC labeling coupled to a "novel" scheme for enriching cysteine containing peptides. The hybrid method has a cysteine enrichment step similar to ICAT. However, ICAT uses complex chemical reagents to modify cysteines in order to enrich cysteine containing peptides. The hybrid method uses much simpler chemistry where cysteines and cystines are oxidized to cysteic acid.

5.1.1. Cysteine Oxidation with Performic Acid

Classic sample preparation methods involve reduction and alkylation of cysteines in order to break the peptide disulfide bonds and prevent the creation of new disulfide bonds. Importantly, these steps ensure that all the cysteines in the sample have the same molecular weight, which facilitates their analysis by MS. The hybrid method also uses cysteine modification for similar reasons, as well as for creating additional negative charges. The complete oxidation of cysteine to cysteic acid, by performic acid treatment, is an irreversible modification that breaks and prevents new disulfide bonds from forming. The treatment of the sample with performic acid is relatively simple and can be performed before or after the digestion step. In the case of in-gel protocols, the peroxide treatment is performed prior to digestion as the immobilization of the proteins in the gel matrix facilitates the removal of the peroxide.

I found that complete oxidation of cysteines and cystines can be achieved with 3% performic acid treatment for 1h at room temperature. Importantly, hydrogen peroxide alone resulted in only partial oxidation of the sample **(Figure 4.2)**. This is most likely due to the inability of hydrogen peroxide to break the disulfide bonds in cystine. Although performic acid can be considered as a strong oxidant, at the concentration used (3%) it

did not artifactually increase the sample complexity, by forming multiple oxidation products of tryptophan, formylation of lysine, oxidation and chlorination of tyrosine etc [229,230]. In addition to the oxidation of cysteine, the only other amino acid that was found to be oxidized by performic acid treatment was methionine. Methionine is typically found partially oxidized in most samples, as this amino acid is very easily oxidized, even by the oxygen in the air.

As already stated, partial oxidation states are generally unwanted as they make the analysis of the sample by mass spectrometry more difficult. For example, if a cysteine is partially or not fully oxidized, it is likely to be missed from the database searches, especially if it is forming new disulfide bonds, as these changes are virtually impossible to account for bioinformatically. The other 3 redox states can be accounted for by allowing for these molecular masses of cysteine, however the addition of these partial modification greatly expand the search space, which has the unwanted effect of increasing the false discovery rate. In addition, partial oxidation states would make the quantitation of a peptide very difficult as it would be split into many forms and all of these forms would have to be taken into account. This would be particularly difficult as these forms would likely have different chromatographic behavior and would elute at different times. These changes in chromatographic behavior would be particularly detrimental for the hybrid technique, which was developed to enrich for cysteic acid residues.

5.1.2. ERLIC Enables Enrichment of Cysteic Acid Peptides

The conversion of cysteine to cysteic acid creates a new strongly negative species in the sample. My initial attempts to enrich for this species focused on ion-exchange chromatography techniques, because ion exchange chromatography depends on the interaction between charged proteins/peptides and an oppositely charged

chromatography medium. In addition, SCX has been used to enrich for phospho-peptides, which also contain a strong negatively charged modification. I explored the use of both SCX and SAX chromatography to enrich cysteic acid peptides. However, I was unable to find conditions where I could efficiently enrich the cysteic acid containing peptides **(Figures 4.4 and 4.5.).** One reason could be that the usage of 'strong' type of resins which refers how much the ionization state of the functional groups vary with pH. This means that a SCX or a SAX has the same charge density on the surface over a broad pH range, whereas the charge density of a weak ion exchanger changes with pH. This limits the selectivity of the strong ion exchangers. I initially tried these resins because of their successful use in enriching for phosphorylated residues. This success is based on the differences in pKa (strength) of phosphoric acid relative to the carboxylic acids found in peptides. However, cysteic acid sits in between phosphoric acid and these carboxylic acids, and perhaps this difference is not great enough to provide enrichment in ion exchange chromatography.

In order to eliminate these factors, I switched to 'weak' type of resins where a peptide with a specific isoelectric point might be selected more easily. Furthermore, a specific chromatography technique based on weak ion exchangers, called ERLIC is particularly useful in providing chromatographic selectivity between similar groups. As my results with SCX and SAX suggested that anion exchange chromatography was superior to cation exchange chromatography **(Figure 4.5)**, I used ERLIC based on the weak anion exchange resin PolyWAX. The enrichment of cysteic acid was successful using ERLIC **(Figure 4.12)**. Moreover, when the conditions are optimized, this enrichment was also easy to perform and could be reduced to three steps that can be performed in a chromatography tip. However, the enrichment is dependent on the pH of the buffers **(Figure 4.8)**, the

loading capacity of the PolyWAX resin **(Figure 4.9),** the amount of starting material **(Figure 4.10)** and the complexity of the sample **(Figure 4.14)**.

5.1.3. Hybrid Method Provides Increased Identification of Peptides and Proteins

The initial optimizations were done using BSA as a model protein, principally because it contains a large number of cysteines and cystines. However, the hybrid method is meant to be used for more complex samples, such as tumor cell lysates. The standard, or classic methods and the hybrid method were therefore compared using whole cell lysates. Approximately 5000 proteins and 13000 peptides were identified with the hybrid method, while the total number of identified proteins and peptides was approximately 1200 and 4200 using the classic method **(Figure 4.15).** These results demonstrate that the hybrid method is useful for getting increased number of protein identifications in complex samples.

This dramatic increase in protein identification is unlikely to be solely due the enhanced enrichment of cysteic acid containing peptides **(Figure 4.14.).** All the fractions produced by the hybrid technique (flow through, wash and elution) are analyzed, and all of them contribute to the increase in protein identification. This is specifically due to the excellent fractionation seen with ERLIC. The high degree of fractionation can be seen by the minimal overlap seen between fractions **(Figures 4.13 and Figure 4.16)**, which means most of the peptides are unique to each fraction and the mass spectrometer is more efficient, because the number of redundant acquisitions is minimized. Furthermore, this makes the quantification of samples subjected to multiple dimension chromatography easier to quantify, as the same peptide signal will not be split over multiple runs.

5.1.4. Hybrid Method Increases the Identification of Post Translational Modifications

When compared to classic techniques, hybrid technique improves the identification of post translationally modified peptides. Database search results indicated that there is also an increased identification in phosphorylation and oxidation sites **(Figure 4.17)**. This shows that hybrid method can also be used for quantitative mapping of in vivo PTMs, especially those that increase the acidity of the peptides such as phosphorylation, oxidation, acetylation, etc.

5.1.5. Hybrid Method is an Accurate and Straightforward Method

The hybrid method was developed with the goal of improving the rate of protein identification and quantification, ideally with minimal sample handling. Although hybrid technique increases the number of identifications, the efficiency, reproducibility and speed of the technique are three other important aspects that make the hybrid technique useful for quantitative differential proteomics. Compared to ICAT, which takes at least 2 days for sample processing, the hybrid technique can be completed in a few hours.

In addition to ICAT, there are many methods where the effects of complex sample preparation is high. In order to avoid variations between samples due to the experimental bias and to minimize sample loss, sample preparation should be minimized to remove unnecessary steps. The hybrid technique can be considered one of the fastest techniques and even with the hybrid technique, unnecessary steps have been removed **(Figure 4.11.)**. Reducing experimental steps is also beneficial when working with many samples and/or biological replicates, since it reduces the total analysis time.

In order to provide increased accuracy, the hybrid technique is coupled to SILAC, a stable isotope labeling approach for accurate quantification of proteins. SILAC uses the metabolic incorporation of stable isotopes in living cells and the quantitative difference in protein levels is then measured by mass-shifts. SILAC allows for the samples to be mixed at a very early stage and this helps minimize "experimenter induced errors" as the mixing ensures that the two (or more) labeled samples are treated identically.

5.1.6. Versatility of the Hybrid Technique

The initial development of the hybrid technique involved coupling the ease of cysteic acid enrichment to the power of SILAC. However, label-free techniques can also be coupled to the enrichment scheme. Label-free techniques are particularly useful when it is impossible or impractical to label the sample metabolically, such as with clinical samples or samples from non-dividing cells, such as neurons.

It is particularly difficult to perform SILAC on non-dividing cells. Since the cultures cannot be expanded, it can be difficult to obtain enough cells to provide the necessary biological replicates, and culture to culture variations may mask any real changes in the experimental conditions. Moreover, SILAC preferentially labels proteins that have high turnover rates. For the proteins that have low or very low turnover rates the stable isotope labeling will not be complete and this will skew the results for these proteins. This is usually not a factor with dividing cells as these unlabeled, stable proteins are diluted out as the culture expands.

In order to test label-free approaches with hybrid method, mouse primary cortical neurons treated with resveratrol and analyzed with hybrid technique. The technique was found successful with label-free approaches, since protein identification is found increased **(Figure 4.20).** Even in workflows where the samples are never mixed, such as in the

various label free techniques, the hybrid method will be advantageous **(Figure 4.20)**, as the increased number of peptide identifications improves the reliability of the quantification. Even in the case of the streamlined hybrid protocol, the standard deviations between biological replicates was higher with the label-free approach than with the SILAC approach.

Another advantage of the hybrid technique is the experimental set up. In these studies, the sample preparation was mostly performed using small columns made from pipette tips. As the columns can be prepared in a few minutes, it is easy to custom make the columns based on the volume or amount of protein in the samples **(Figure 4.11)**.

Although, I typically performed the purification steps manually, throughput of the technique can be improved by integrating the columns into 1.5 mL tubes and the purification steps can be performed with the aid of a centrifuge [231].

5.2. Limitations of the Hybrid Technique

The limitations of hybrid technique mainly come from the isotopic labeling technique itself. As mentioned before, while providing an accurate quantification, SILAC increasing the sample complexity. In isotope labeling strategies, the stable isotope tags produce a mass shift, which is used to track which experimental group (sample) produced that peptide. In this case of a binary system, each peptide appears as two peaks, one from the heavy labeled sample and one from the light. This effectively doubles the sample complexity and is much worse for higher orders of multiplexing. This isotope-induced increase in sample complexity is one of the main advantages of cysteic acid enrichment. As cysteine is found in about 5% of peptides, the enrichment will immediately give a 20 fold loss of complexity.

Sample complexity can also limit the efficiency of the hybrid technique. While the hybrid technique works well when the complexity is low, the efficiency is lower when the complexity is high. However, this is compensated by the fact that there is very little overlap between the three fractions produced by the hybrid technique.

5.3. Investigation of Sirtinol-Induced Senescence in Tumor Cells by Hybrid Method

In order to verify that the hybrid technique can be used on "real world" samples, I used it to investigate the response of tumor cells to therapeutic compounds, such as sirtinol. Sirtinol has been found to induce senescence in tumor cells and is being explored as potential therapeutic agent. In fact, the role of senescence in tumor progression has been shown in many studies. Conventional anticancer therapies have also indicated that senescence plays a role, even when cell death is the main effect of the treatment [232–234].

Driving tumor cells into senescence could have more benefits than simply killing the cells. In contrast to apoptosis, senescence may have fewer side effects related to toxicity [235] and senescent cells also change the microenvironment, including the recruitment of immune cells, which may help eliminate any actively proliferating tumor cells. Therefore, there is growing interest in using chemically induced senescence as a treatment for some cancers. With that idea in mind, I wanted to explore the effects of sirtinol, an HDAC inhibitor that has been used to therapeutically induce senescence in some tumor cells. I believed this could be an ideal system to test the hybrid method, as quantitative proteomics is difficult to perform on a dying cell population.

5.3.1. Sirtinol Induces Senescence in H1299 Tumor Cells

I initially chose H1299 cells as they were already being used in the laboratory and the first step was to determine if sirtinol could induce senescence in these cells. I was able to show that a single, high dose of sirtinol was enough to cause a long term (10 days) block in cell proliferation **(Figure 4.22.).** This was also accompanied by a morphological change in the cells, which became large and flat **(Figure 4.23 A, B and C)**.

Senescence is typically induced and maintained by the tumor suppressor Rb. Rb is found to be hypo-phosphorylated in senescent cells and the hypo-phosphorylated form is also the form that actively suppresses the E2F transcription factor. Sirtinol treated H1299 cells were checked for Rb phosphorylation, and hypo-phosphorylation was observed on three different points on Rb protein **(Figure 4.31).** These data indicate that, similar to other senescent cells, Rb plays a role in inhibiting cell proliferation in sirtinol treated cells.

The block in cell proliferation and the morphological change are indicative of sirtinol inducing senescence in the H1299 cells. However, to verify that senescence is being induced, I stained the cells for a histochemical marker of senescence, SA-β-gal activity [104]. In fact, sirtinol caused an increase in the histochemical staining for SA-β-gal activity **(Figure 4.23.)**. In order to verify these in-situ experiments I checked the quantitative MS experiments for the relative levels of the β-galactosidase responsible for this staining. Indeed, there was a 1.7 fold increase for this β-galactosidase between control and 300 uM sirtinol treated samples. This indicates that the SA-β-gal activity detected by X-gal in response to sirtinol, is due to the induction of the same protein induced by other forms of senescence. Taken together, these data indicate that sirtinol induces senescence in H1299 cells.

Interestingly, there was a difference in the behavior of the cells when treated with a lower does (200 uM) of sirtinol. A single 300 uM sirtinol treatment caused a long term inhibition of cell proliferation and resulted in ~90% of the cells becoming SA-β-gal positive. Whereas, 200 um sirtinol was only able to delay the initiation of cell proliferation for a few days and resulted in far fewer SA-β-gal positive cells. Cell proliferation was probably reinitiated in the 200 uM treated sample by these cells that at least by SA-β-gal did not undergo senescence. Both doses of sirtinol are well above the dose required to inhibit Sirt1, so it is unclear why such high doses are necessary. This could be due to cell specific factors, such as the presence of multidrug transporters in the plasma membrane, or it may indicate that Sirt1 is not the only sirtinol target. Furthermore, I found that two doses of 200 uM sirtinol spaced 1 day apart also resulted in a long term inhibition of cell proliferation (data not shown), which suggests that sirtinol may be more effective at certain stages of the cell cycle.

5.3.2. Quantitative Analysis of Sirtinol Treated Tumor Cells

In order to better understand the effects of sirtinol on H1299 cells, I used the hybrid method to quantitate the changes produced by sirtinol at the proteome level. I used SILAC labeling for ten generations and then one pool of cells was treated for 1 day with 300 uM sirtinol. Cells were harvested 8 days after sirtinol treatment and quantitative data was obtained using the hybrid method coupled to mass spectrometry.

This workflow resulted in approximately 4000 quantified proteins and 228 of these proteins were found to be changed by more than two-fold. Specifically, 88 proteins were up regulated and 140 proteins were down regulated **(Table 4.1 and Table 4.2)**. As expected, most of the protein levels were not changed. Importantly, this indicates that the samples were properly mixed. For example if the twice as much of one sample was

added, then most proteins would show a 2 fold change in expression levels. This shows that the hybrid technique can be used for binary comparisons such as characterizing the response of mammalian cells to pharmacological agents.

In order to further validate the results obtained by the hybrid method, the levels of two downregulated proteins (FAS and AHNAK) **(Figure 4.27)** and two upregulated proteins (EIF2α and PPP2R3A) **(Figure 4.28)** were analyzed by western blotting. Although these blots were not quantitated, they do indicate that the hybrid technique can identify both up and down regulated proteins. It is particularly difficult to quantitate protein levels accurately by western blotting and in order to verify that the absolute change in levels reported by the mass spectrometry is correct would require a truly quantitative technique, such as ELISA or a targeted mass spectrometry based approach, such as multiple reaction monitoring (MRM).

5.3.3. Biological Relevance of Altered Proteins

The functional annotations of a protein can be used to help decipher which proteins are critical for the cells response to sirtinol. This is particularly useful as the quantitative MS results have already reduced the number of candidates down to ~200. However, even on a short list of 200 proteins this can be daunting task and probably cannot be done manually. For example, I further characterized FAS protein levels in order to validate the MS results. However, FAS was not chosen randomly.

Fatty acid metabolic pathways have been previously reported to be associated with carcinogenesis and FAS can also be used as a biomarker for certain cancers [236]. In fact, the FAS antibodies were already being used in the laboratory as a marker for liver cancer. Fatty acid synthase (FAS) is an important enzyme which involved in endogenous lipogenesis in mammals and also responsible for catalyzing the synthesis of long-chain fatty

acids. FAS has also been identified as a crucial factor for sustaining several biological features of cancer cells [237].

In all sirtinol treated tumor cells, FAS levels were found downregulated. I checked whether this situation has any relevance with senescence and cancer states. Indeed, FAS levels has been previously found at high levels in a variety of tumors [238–240] while in normal tissues it remains at normal levels. Also, it has been reported in several studies that inhibition of FAS expression suppresses cancer cell proliferation *in vivo* and *in vitro* [241,242]. Hence, a reduction in FAS levels may be one of the therapeutic effects of sirtinol and inhibition of FAS has even been proposed as a potential therapeutic strategy [243].

If FAS is playing a role in sirtinol induced senescence, then it would also be expected to be important in other forms of senescence, such as replicative senescence. In order to check the levels of FAS during replicative senescence, I used MEF cells. MEF cells were cultured until they underwent crisis, which is a form of senescence, and then upon further passaging a few cells escape crisis and begin to proliferate and they quickly take over the culture. Interestingly, there was a correlation between FAS levels and replication of the cells **(Figure 4.33)**. FAS levels began decreasing at the start of replicative crisis and remained low until the cell proliferation resumed. These results further highlight the possible role of FAS in senescence, cancer and aging [244]. However, it is possible that FAS is simply a marker of a high metabolic rate and can be thought of as a metabolic marker of cell proliferation [245].

FAS was chosen for further study, mostly because my laboratory was already familiar with its role in cancer. However, one of the strengths of the proteomics approach is being able to uncover pathways that were not previously known to be key players in senescence and bioinformatics approaches are used to uncover these relationships.

5.4. Bioinformatics

High throughput expression level data, both transcriptomic and proteomic, can be difficult to interpret, because these approaches typically uncover a large number of changes. However, layering this expression data on top of data such as metabolic pathways, interaction maps, or even Gene Ontologies (GO) is often used to help discover common functional pathways that are part of the cellular response. Many of these changes will be part of the biological response and many will not. Pathway analysis is a powerful way to identify the proteins involved in the response, as typically many proteins in a pathway will be altered.

STRING is a databases of known and predicted protein: protein interactions and it can be used to help uncover related proteins [246]. Protein interaction maps are a particularly powerful way to analyze proteomics data, because most proteins work in concert with other proteins. In fact, early work on the high throughput mapping of protein interactions was justified because mapping protein-protein interactions is a quick way to functionally annotate genomic databases.

In order to help uncover new relationships among the proteins whose expression level changes in response to sirtinol, I analyzed both the up and down regulated proteins using the STRING database **(Figure 5.1. and 5.2.)**. Importantly STRING uncovered several pathways that were strongly enriched in either the up or down regulated protein lists. Significantly, protein-protein interactions involved in regulating protein synthesis were found in the upregulated protein lists and include members of the mTOR signaling cascade, including the translation elongation and initiation factors (EIF proteins) **(Figure 5.1.)**. Significantly, the mTOR pathway has been shown to be responsible for regulating the translation of SASP in senescent cells [247]. Specifically when the mTOR complex is

inhibited with rapamycin, SASP factors were also found to be reduced. Hence, up-regulation of mTOR activity by sirtinol is expected to regulate SASP in H1299 cells after treatment.

Eukaryotic initiation factor proteins (EIF2, EIF3 proteins) and ribosomal proteins (RL8, RL17, and RL23) also clustered in the protein interaction map shown in Figure 5.1. Up-regulation of these proteins may indicate a change in protein synthesis that promotes the senescent phenotype and maybe regulated by mTOR, as these proteins were also found interacting with mTOR (**Figure 5.1**).

Another cluster of interactions is centered on the histone proteins. The histones are the known targets of the sirtuin family of histone deacetylases. One of the most interesting proteins in the histone protein cluster is H2AX (histone H2A variant) which is a well-known biomarker for DNA-damage (specifically DSBs) [248]. H2AX has been found to be up-regulated and phosphorylated in senescent cells [248]. The sirtinol dependent upregulation of histone proteins is consistent with the chromatin remodeling and DNA-damage markers seen in other models of senescence. Furthermore, these results suggest that sirtinol is inducing senescence via well-known senescence pathways.

One of the weaknesses of this type of network analysis is that it does not provide additional information about those proteins such as PPR23A, FLNC and WNK2 that do no connect with other proteins found to be up-regulated (**Figure 5.1**). This may be due to lack of experimental evidence or their activities are not really required for sirtinol induced senescence.

Similarly, STRING also revealed the presence of a single protein-protein interaction network in the down regulated proteins **(Figure 5.2)**. Many of these down regulated proteins are involved in regulating protein quality. Including chaperones (BAG3), the

ubiquitination machinery (UCHL1, UBA1, RN213, PFWD3, HECW2, HERC1, TRIM25), proteasome subunits (PSA6), prolyl isomerases (PPIA). It is currently unclear why the protein quality control machinery is downregulated during senescence, but this may be related to the upregulation of mTOR activity, as mTOR also functions to inhibit the protein quality control system, and specifically BAG3 and the proteasome [249].

As with the up-regulated proteins, many down-regulated proteins did not form connections, such as AHNAK. AHNAK's function and its involvement in several states such as cancer has been recently investigated [250]. AHNAK is involved in different cellular processes and pathways, ranging from regulating calcium channels, cell architecture, and cell migration. AHNAK has been shown to be highly abundant in lipid rafts of breast cancer cells and AHNAK is suggested to be involved in tumor metastasis [250]. However there are also studies suggesting that AHNAK may work as a tumor suppressor[251]. Although AHNAK has several different functions, it was not found to interact with other proteins downregulated by sirtinol. AHNAK's function in response to sirtinol needs further characterization.

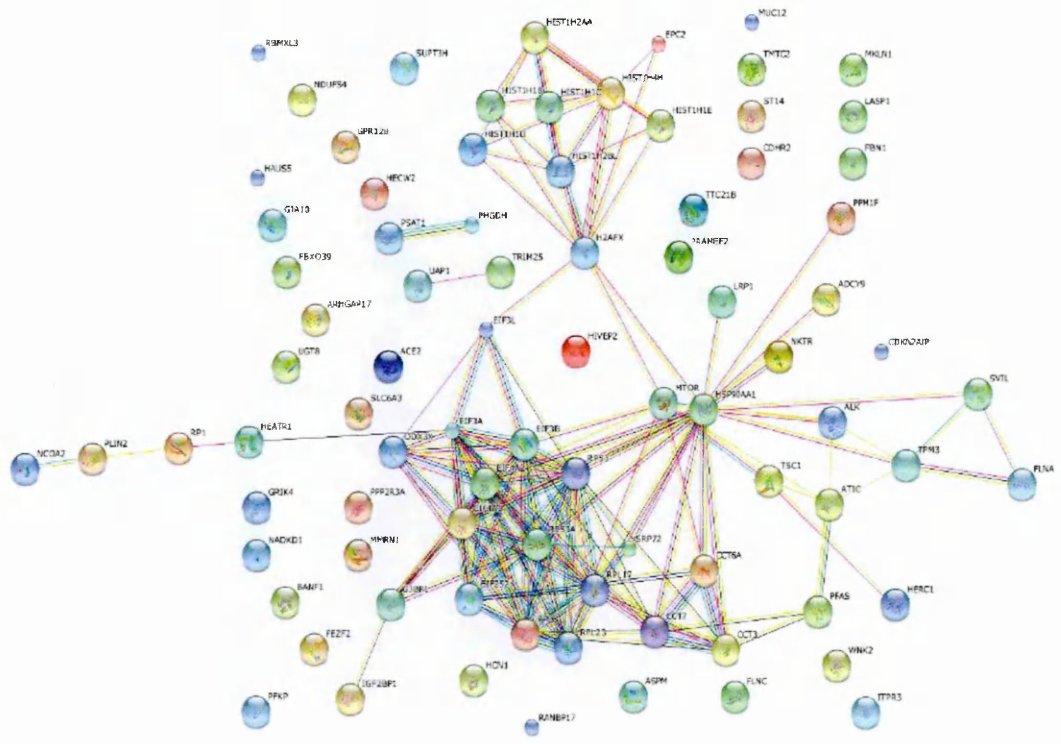


Figure 5.1. Protein interaction mapping of sirtinol up-regulated proteins. The list of proteins up-regulated by sirtinol was analyzed using STRING [246]. Protein: protein interactions found in the database are indicated by connections being drawn between the proteins.

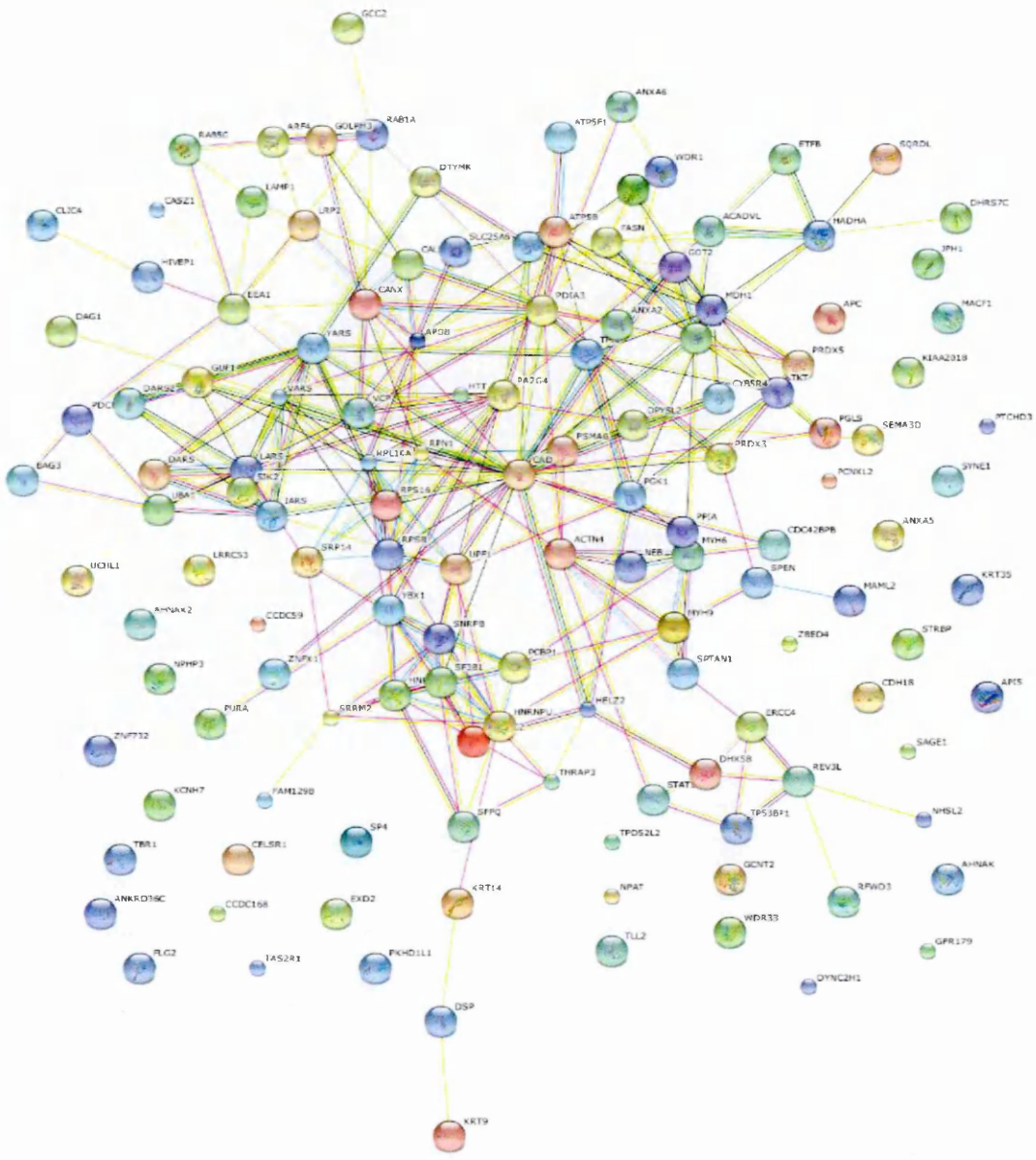


Figure 5.2. Protein interaction mapping of sirtinol down-regulated proteins. The list of proteins down-regulated by sirtinol was analyzed using STRING [246]. Protein: protein interactions found in the database are indicated by connections being drawn between the proteins.

5.5. Future Work

The Hybrid method was developed with the aim of maximizing protein identifications and it was applied to determine how a tumor cell line responds to the HDAC inhibitor sirtinol.

During the experiments I used a gel-based separation before cysteic acid enrichment in order to reduce the complexity of the samples, specifically for cell lysates. I believe that with more experimentation the gel separation and in-gel digestion steps can actually be removed from the method and cysteine enrichment can then be performed on total lysate. Protein identification from total cell lysate has been previously shown with a relatively simple technique based on precipitation of the samples onto a chromatographic bed packed in a pipette tip [231]. I believe that this technique can be optimized to be used in the cysteic acid enrichment process. Although conceptually simple several optimizations should be considered such as the removing any contaminants, such as detergents that might degrade the chromatographic purification.

The quantitative mass spectrometry analyses produced a list of differentially expressed proteins and I have begun confirming the changes in protein levels using another assay. But so far the expression levels for only a few of them have been verified by western blotting. This is necessary to determine if the changes reported by the hybrid method are reliable. Ideally the levels of all the proteins should be verified by another method, however this is not possible, and a more through bioinformatics analysis should be performed to focus the verification process onto those proteins that are most likely to be involved in senescence.

Importantly, the relevance of the changes in protein levels reported by the hybrid method also needs to be studied in a biological context. Especially in their roles in establishing and maintaining the senescent phenotype in response to sirtinol. This is best performed with

gene knockdowns using siRNA and over expression studies. For example, in the case of FAS, siRNA knockdown may result in a lower dose of sirtinol being needed to induce the senescent phenotype, and the knockdown cells may even become senescent without chemical treatment. Over expression of FAS, on the other hand, would be expected to make it even more difficult for sirtinol to produce a fully senescent culture. These cells with altered protein levels can also be subjected to another round of quantitative mass spectrometry to determine, if any of the other changes in protein level are dependent on FAS. Moreover, that the proteins that interact with FAS can be identified, because these proteins are also likely to be involved in regulating the transition to the senescent state.

My results also revealed an increase in the number of acetylated proteins in response to sirtinol treatment. This was not surprising given that sirtinol is an HDAC inhibitor. However, these acetylated proteins were found many days after the cessation of sirtinol treatment. Therefore, I believe that characterizing the "acetylome", as many acetylated proteins as possible, would also reveal interesting results. This would be especially powerful, if it were performed at different time points, starting immediately after sirtinol treatment. Although immunological methods could be used to enrich for acetylated peptides, I believe that ERLIC chromatography could also be adapted to these studies as acetylation of lysine residues changes its charge state. This would likely require a new round of optimization steps. However, if successful, would make ERLIC a very important technique for proteomic studies.

Tremendous efforts have been put into cancer research, and there is a critical need for the discovery of new therapeutic compounds. Quantitative proteomics can also be used to characterize the direct mechanism of action of these drugs. Most drugs have more than one target, these are the targets for which the drug was developed, Sirt1, plus all the other "non-specific" targets that also interact with the sirtinol. Proteomic techniques

can be used to identify both the specific and non-specific actions of a drug, like sirtinol [252, 234]. For example, sirtinol-coated beads can be used as an affinity reagent to perform pull-down experiments with the interacting proteins being identified by mass spectrometry. Although sirtinol is a known inhibitor of Sirt I, the full spectrum of interacting proteins can help reveal any unknown targets of the compound. This is especially interesting, for a drug like sirtinol where the doses required to induce senescence are well above the doses required to inhibit the known target.

The hybrid technique that I developed can be adapted to virtually any proteomics workflow and not just those requiring cysteine enrichment to reduce sample complexity. I used the technique specifically to characterize how sirtinol affects a specific tumor cell line and the continuation of these studies should result in a better understanding of sirtinol action, as well as, expanding the use of the hybrid method I developed.

6. REFERENCES

1. Chandramouli, K. & Qian, P.-Y. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum. Genomics Proteomics* **2009,** (2009).

2. Mitchell, P. Proteomics retrenches. *Nat. Biotechnol.* **28,** 665–70 (2010).

3. Hall, J. A., Dominy, J. E., Lee, Y. & Puigserver, P. The sirtuin family's role in aging and age-associated pathologies. *J. Clin. Invest.* **123,** 973–9 (2013).

4. Tyers, M. & Mann, M. From genomics to proteomics. *Nature* **422,** 193–7 (2003).

5. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13,** 227–32 (2012).

6. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19,** 1720–30 (1999).

7. Wilkins, M. R. *et al.* From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology. (N. Y).* **14,** 61–5 (1996).

8. Lodish, H. *et al.* Molecular Cell Biology. 5th Edition. Section 3.2. Folding, Modification, and Degradation of Proteins. 86–91 (2003).

9. Breitbart, R. E., Andreadis, A. & Nadal-Ginard, B. Alternative splicing: a ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annu. Rev. Biochem.* **56,** 467–95 (1987).

10. Mann, M. & Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21,** 255–61 (2003).

11. Tan, S., Tan, H. T. & Chung, M. C. M. Membrane proteins and membrane proteomics. *Proteomics* **8,** 3924–32 (2008).

12. Siuti, N. & Kelleher, N. L. Decoding protein modifications using top-down mass spectrometry. *Nat. Methods* **4,** 817–21 (2007).

13. Bodzon-Kulakowska, A. *et al.* Methods for samples preparation in proteomic research. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **849,** 1–31 (2007).

14. Zhou, J.-Y. *et al.* Simple sodium dodecyl sulfate-assisted sample preparation method for LC-MS-based proteomics applications. *Anal. Chem.* **84,** 2862–7 (2012).

15. Thakur, D. *et al.* Microproteomic analysis of 10,000 laser captured microdissected breast tumor cells using short-range sodium dodecyl sulfate-polyacrylamide gel

electrophoresis and porous layer open tubular liquid chromatography tandem mass spectrometry. *J. Chromatogr. A* **1218,** 8168–74 (2011).

16. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1,** 2856–60 (2006).

17. Granvogl, B., Plöscher, M. & Eichacker, L. A. Sample preparation by in-gel digestion for mass spectrometry-based proteomics. *Anal. Bioanal. Chem.* **389,** 991–1002 (2007).

18. Olsen, J. V, Ong, S.-E. & Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **3,** 608–14 (2004).

19. Protein Digestion - Creative Proteomics. at <http://www.creative-proteomics.com/Services/Digestion-in-gel-or-in-solution.htm>

20. Finishing the euchromatic sequence of the human genome. *Nature* **431,** 931–45 (2004).

21. Jensen, O. N. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem. Biol.* **8,** 33–41 (2004).

22. Anderson, N. L. & Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1,** 845–67 (2002).

23. Rabilloud, T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* **2,** 3–10 (2002).

24. Hortin, G. L., Jortani, S. A., Ritchie, J. C., Valdes, R. & Chan, D. W. Proteomics: a new diagnostic frontier. *Clin. Chem.* **52,** 1218–22 (2006).

25. Niessen, S., McLeod, I. & Yates, J. R. HPLC separation of digested proteins and preparation for matrix-assisted laser desorption/ionization analysis. *CSH Protoc.* **2006,** (2006).

26. Molnár, I. & Horváth, C. Reverse-phase chromatography of polar biological substances: separation of catechol compounds by high-performance liquid chromatography. *Clin. Chem.* **22,** 1497–502 (1976).

27. Alpert, A. J. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal. Chem.* **80,** 62–76 (2008).

28. Di Palma, S., Hennrich, M. L., Heck, A. J. R. & Mohammed, S. Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. *J. Proteomics* **75,** 3791–813 (2012).

29. Delahunty, C. & Yates, J. R. Protein identification using 2D-LC-MS/MS. *Methods* **35,** 248–55 (2005).

30. Hörth, P., Miller, C. A., Preckel, T. & Wenz, C. Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Mol. Cell. Proteomics* **5,** 1968–74 (2006).

31. Bonifacino, J. S., Dell'Angelica, E. C. & Springer, T. A. Immunoprecipitation. *Curr. Protoc. Mol. Biol.* **Chapter 10,** Unit 10.16 (2001).

32. Patel, B. B. *et al.* Assessment of two immunodepletion methods: off-target effects and variations in immunodepletion efficiency may confound plasma proteomics. *J. Proteome Res.* **11,** 5947–58 (2012).

33. Bellei, E. *et al.* High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids* **40,** 145–56 (2011).

34. Righetti, P. G., Castagna, A., Antonioli, P. & Boschetti, E. Prefractionation techniques in proteome analysis: the mining tools of the third millennium. *Electrophoresis* **26,** 297–319 (2005).

35. Selvaraju, S. & Rassi, Z. El. Liquid-phase-based separation systems for depletion, prefractionation and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis--an update covering the period 2008-2011. *Electrophoresis* **33,** 74–88 (2012).

36. Lahm, H. W. & Langen, H. Mass spectrometry: a tool for the identification of proteins separated by gels. *Electrophoresis* **21,** 2105–14 (2000).

37. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422,** 198–207 (2003).

38. Nesvizhskii, A. I. Protein identification by tandem mass spectrometry and sequence database searching. *Methods Mol. Biol.* **367,** 87–119 (2007).

39. Walther, T. C. & Mann, M. Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.* **190,** 491–500 (2010).

40. Todd, J. F. J. Recommendations for nomenclature and symbolism for mass spectroscopy. *Int. J. Mass Spectrom. Ion Process.* **142,** 209–240 (1995).

41. Nesvizhskii, A. I., Vitek, O. & Aebersold, R. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat. Methods* **4,** 787–97 (2007).

42. Dancík, V., Addona, T. A., Clauser, K. R., Vath, J. E. & Pevzner, P. A. De novo peptide sequencing via tandem mass spectrometry. *J. Comput. Biol.* **6,** 327–42

43. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–89 (1994).

44. Mann, M. & Wilm, M. Error-Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Anal. Chem.* **66,** 4390–4399 (1994).

45. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–67 (1999).

46. Craig, R. & Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20,** 1466–7 (2004).

47. Geer, L. Y. *et al.* Open mass spectrometry search algorithm. *J. Proteome Res.* **3,** 958–64

48. Fenyö, D. & Beavis, R. C. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Anal. Chem.* **75,** 768–74 (2003).

49. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–58 (2003).

50. Huang, T., Wang, J., Yu, W. & He, Z. Protein inference: a review. *Brief. Bioinform.* **13,** 586–614 (2012).

51. Rappsilber, J. & Mann, M. What does it mean to identify a protein in proteomics? *Trends Biochem. Sci.* **27,** 74–8 (2002).

52. Han, D. K., Eng, J., Zhou, H. & Aebersold, R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19,** 946–51 (2001).

53. ProteolQ:: High-throughput qualitative and quantitative analysis software suite. at <http://www.premierbiosoft.com/protein_quantification_software/index.html>

54. Kumar, C. & Mann, M. Bioinformatics analysis of mass spectrometry-based proteomics data sets. *FEBS Lett.* **583,** 1703–12 (2009).

55. Smithies, O. How it all began: a personal history of gel electrophoresis. *Methods Mol. Biol.* **869,** 1–21 (2012).

56. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–5 (1970).

57. Shapiro, A. L., Viñuela, E. & Maizel, J. V. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem. Biophys. Res. Commun.* **28,** 815–20 (1967).

58. Görg, A., Weiss, W. & Dunn, M. J. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **4,** 3665–85 (2004).

59. Gauci, V. J., Wright, E. P. & Coorssen, J. R. Quantitative proteomics: assessing the spectrum of in-gel protein detection methods. *J. Chem. Biol.* **4,** 3–29 (2011).

60. Lilley, K. S. & Friedman, D. B. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev. Proteomics* **1,** 401–9 (2004).

61. Voyksner, R. D. & Lee, H. Investigating the use of an octupole ion guide for ion storage and high-pass mass filtering to improve the quantitative performance of electrospray ion trap mass spectrometry. *Rapid Commun. Mass Spectrom.* **13,** 1427–37 (1999).

62. Zhu, W., Smith, J. W. & Huang, C.-M. Mass spectrometry-based label-free quantitative proteomics. *J. Biomed. Biotechnol.* **2010,** 840518 (2010).

63. Wang, M., You, J., Bemis, K. G., Tegeler, T. J. & Brown, D. P. G. Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. *Brief. Funct. Genomic. Proteomic.* **7,** 329–39 (2008).

64. Lundgren, D. H., Hwang, S.-I., Wu, L. & Han, D. K. Role of spectral counting in quantitative proteomics. *Expert Rev. Proteomics* **7,** 39–53 (2010).

65. Deracinois, B., Flahaut, C., Duban-Deweer, S. & Karamanos, Y. Comparative and Quantitative Global Proteomics Approaches: An Overview. *Proteomes* **1,** 180–218 (2013).

66. Schmidt, A., Kellermann, J. & Lottspeich, F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5,** 4–15 (2005).

67. Ross, P. L. *et al.* Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3,** 1154–69 (2004).

68. Gygi, S. P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–9 (1999).

69. Goshe, M. B. *et al.* Phosphoprotein Isotope-Coded Affinity Tag Approach for Isolating and Quantitating Phosphopeptides in Proteome-Wide Analyses. *Anal. Chem.* **73,** 2578–2586 (2001).

70. Bottari, P., Aebersold, R., Turecek, F. & Gelb, M. H. Design and synthesis of visible isotope-coded affinity tags for the absolute quantification of specific proteins in complex mixtures. *Bioconjug. Chem.* **15,** 380–8

71. Yi, E. C. *et al.* Increased quantitative proteome coverage with (13)C/(12)C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomics* **5,** 380–7 (2005).

72. Sevinsky, J. R., Brown, K. J., Cargile, B. J., Bundy, J. L. & Stephenson, J. L. Minimizing back exchange in 18O/16O quantitative proteomics experiments by

incorporation of immobilized trypsin into the initial digestion step. *Anal. Chem.* **79,** 2158–62 (2007).

73. Perez, J. F. & Reeds, P. J. A new stable isotope method enables the simultaneous measurement of nucleic acid and protein synthesis in vivo in mice. *J. Nutr.* **128,** 1562–9 (1998).

74. Beynon, R. J. & Pratt, J. M. Metabolic labeling of proteins for proteomics. *Mol. Cell. Proteomics* **4,** 857–72 (2005).

75. Ong, S.-E. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* **1,** 376–386 (2002).

76. Ong, S.-E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–86 (2002).

77. Ong, S.-E., Kratchmarova, I. & Mann, M. Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2,** 173–81

78. National Institute on Aging - Proteomics. at <http://www.grc.nia.nih.gov/branches/lci/muproteomics.htm>

79. Weckwerth, W., Willmitzer, L. & Fiehn, O. Comparative quantification and identification of phosphoproteins using stable isotope labeling and liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.* **14,** 1677–81 (2000).

80. Ong, S.-E., Mittler, G. & Mann, M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods* **1,** 119–26 (2004).

81. Bonenfant, D. *et al.* Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. *Mol. Cell. Proteomics* **6,** 1917–32 (2007).

82. Vermeulen, M. *et al.* Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131,** 58–69 (2007).

83. Sol, E. M. *et al.* Proteomic investigations of lysine acetylation identify diverse substrates of mitochondrial deacetylase sirt3. *PLoS One* **7,** e50545 (2012).

84. Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–40 (2009).

85. Weinert, B. T. *et al.* Proteome-wide mapping of the Drosophila acetylome demonstrates a high degree of conservation of lysine acetylation. *Sci. Signal.* **4,** ra48 (2011).

86. Udeshi, N. D., Mertins, P., Svinkina, T. & Carr, S. A. Large-scale identification of ubiquitination sites by mass spectrometry. *Nat. Protoc.* **8,** 1950–60 (2013).

87. Meierhofer, D., Wang, X., Huang, L. & Kaiser, P. Quantitative analysis of global ubiquitination in HeLa cells by mass spectrometry. *J. Proteome Res.* **7,** 4566–76 (2008).

88. Fíla, J. & Honys, D. Enrichment techniques employed in phosphoproteomics. *Amino Acids* **43,** 1025–47 (2012).

89. Wu, C. C. & MacCoss, M. J. Quantitative proteomic analysis of mammalian organisms using metabolically labeled tissues. *Methods Mol. Biol.* **359,** 191–201 (2007).

90. McClatchy, D. B., Dong, M.-Q., Wu, C. C., Venable, J. D. & Yates, J. R. 15N metabolic labeling of mammalian tissue with slow protein turnover. *J. Proteome Res.* **6,** 2005–10 (2007).

91. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R. & Mann, M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* **7,** 383–385 (2010).

92. Ishihama, Y. *et al.* Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat. Biotechnol.* **23,** 617–21 (2005).

93. Molina, H. *et al.* Temporal profiling of the adipocyte proteome during differentiation using a five-plex SILAC based strategy. *J. Proteome Res.* **8,** 48–58 (2009).

94. Han, Z. *et al.* Role of p21 in apoptosis and senescence of human colon cancer cells treated with camptothecin. *J. Biol. Chem.* **277,** 17154–60 (2002).

95. Schmitt, C. A. *et al.* A senescence program controlled by p53 and p16INK4a contributes to the outcome of cancer therapy. *Cell* **109,** 335–46 (2002).

96. Te Poele, R. H., Okorokov, A. L., Jardine, L., Cummings, J. & Joel, S. P. DNA damage is able to induce senescence in tumor cells in vitro and in vivo. *Cancer Res.* **62,** 1876–83 (2002).

97. Shay, J. W. & Roninson, I. B. Hallmarks of senescence in carcinogenesis and cancer therapy. *Oncogene* **23,** 2919–33 (2004).

98. Elmore, L. W. *et al.* Adriamycin-induced senescence in breast tumor cells involves functional p53 and telomere dysfunction. *J. Biol. Chem.* **277,** 35509–15 (2002).

99. Roninson, I. B. Tumor senescence as a determinant of drug response in vivo. *Drug Resist. Updat.* **5,** 204–8 (2002).

100. Haigis, M. C. & Guarente, L. P. Mammalian sirtuins--emerging roles in physiology, aging, and calorie restriction. *Genes Dev.* **20,** 2913–21 (2006).

101. Tan, J., Cang, S., Ma, Y., Petrillo, R. L. & Liu, D. Novel histone deacetylase inhibitors in clinical trials as anti-cancer agents. *J. Hematol. Oncol.* **3,** 5 (2010).

102. Hayflick, L. The limited in vitro lifetime of human diploid cell strains. *Exp. Cell Res.* **37,** 614–636 (1965).

103. Campisi, J. Cellular senescence as a tumor-suppressor mechanism. *Trends Cell Biol.* **11,** S27–31 (2001).

104. Dimri, G. P. *et al.* A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 9363–7 (1995).

105. Lee, B. Y. *et al.* Senescence-associated beta-galactosidase is lysosomal beta-galactosidase. *Aging Cell* **5,** 187–95 (2006).

106. Green, D. R. & Evan, G. I. A matter of life and death. *Cancer Cell* **1,** 19–30 (2002).

107. Chen, Q. M., Liu, J. & Merrett, J. B. Apoptosis or senescence-like growth arrest: influence of cell-cycle position, p53, p21 and bax in H2O2 response of normal human fibroblasts. *Biochem. J.* **347,** 543–51 (2000).

108. Tepper, C. G., Seldin, M. F. & Mudryj, M. Fas-mediated apoptosis of proliferating, transiently growth-arrested, and senescent normal human fibroblasts. *Exp. Cell Res.* **260,** 9–19 (2000).

109. Zhang, H., Pan, K.-H. & Cohen, S. N. Senescence-specific gene expression fingerprints reveal cell-type-dependent physical clustering of up-regulated chromosomal loci. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 3251–6 (2003).

110. Ressler, S. *et al.* p16INK4A is a robust in vivo biomarker of cellular aging in human skin. *Aging Cell* **5,** 379–89 (2006).

111. Coppé, J.-P., Desprez, P.-Y., Krtolica, A. & Campisi, J. The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu. Rev. Pathol.* **5,** 99–118 (2010).

112. Takai, H., Smogorzewska, A. & de Lange, T. DNA damage foci at dysfunctional telomeres. *Curr. Biol.* **13,** 1549–56 (2003).

113. d'Adda di Fagagna, F. *et al.* A DNA damage checkpoint response in telomere-initiated senescence. *Nature* **426,** 194–8 (2003).

114. Herbig, U., Jobling, W. A., Chen, B. P. C., Chen, D. J. & Sedivy, J. M. Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). *Mol. Cell* **14,** 501–13 (2004).

115. Campisi, J., Andersen, J. K., Kapahi, P. & Melov, S. Cellular senescence: a link between cancer and age-related degenerative disease? *Semin. Cancer Biol.* **21,** 354–9 (2011).

116. Freund, A., Orjalo, A. V, Desprez, P.-Y. & Campisi, J. Inflammatory networks during cellular senescence: causes and consequences. *Trends Mol. Med.* **16,** 238–46 (2010).

117. Liu, H. *et al.* Augmented Wnt signaling in a mammalian model of accelerated aging. *Science* **317,** 803–6 (2007).

118. Chung, H. Y. *et al.* Molecular inflammation: underpinnings of aging and age-related diseases. *Ageing Res. Rev.* **8,** 18–30 (2009).

119. Campisi, J. Aging, cellular senescence, and cancer. *Annu. Rev. Physiol.* **75,** 685–705 (2013).

120. Rodier, F. *et al.* Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion. *Nat. Cell Biol.* **11,** 973–9 (2009).

121. Pazolli, E. *et al.* Chromatin remodeling underlies the senescence-associated secretory phenotype of tumor stromal fibroblasts that supports cancer progression. *Cancer Res.* **72,** 2251–61 (2012).

122. Shay, J. W. & Wright, W. E. Hayflick, his limit, and cellular ageing. *Nat. Rev. Mol. Cell Biol.* **1,** 72–6 (2000).

123. Harley, C. B., Futcher, A. B. & Greider, C. W. Telomeres shorten during ageing of human fibroblasts. *Nature* **345,** 458–60 (1990).

124. Wright, W. E., Piatyszek, M. A., Rainey, W. E., Byrd, W. & Shay, J. W. Telomerase activity in human germline and embryonic tissues and cells. *Dev. Genet.* **18,** 173–9 (1996).

125. d'Adda di Fagagna, F., Teo, S.-H. & Jackson, S. P. Functional links between telomeres and proteins of the DNA-damage response. *Genes Dev.* **18,** 1781–99 (2004).

126. Martens, U. M., Chavez, E. A., Poon, S. S., Schmoor, C. & Lansdorp, P. M. Accumulation of short telomeres in human fibroblasts prior to replicative senescence. *Exp. Cell Res.* **256,** 291–9 (2000).

127. Hemann, M. T., Strong, M. A., Hao, L. Y. & Greider, C. W. The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell* **107,** 67–77 (2001).

128. Unal, E. *et al.* DNA damage response pathway uses histone modification to assemble a double-strand break-specific cohesin domain. *Mol. Cell* **16,** 991–1002 (2004).

129. McEachern, M. J., Krauskopf, A. & Blackburn, E. H. Telomeres and their control. *Annu. Rev. Genet.* **34,** 331–358 (2000).

130. Weng, N. P. & Hodes, R. J. The role of telomerase expression and telomere length maintenance in human and mouse. *J. Clin. Immunol.* **20,** 257–67 (2000).

131. Wright, W. E. & Shay, J. W. Telomere dynamics in cancer progression and prevention: fundamental differences in human and mouse telomere biology. *Nat. Med.* **6,** 849–51 (2000).

132. Zeng, X. & Rao, M. S. Human embryonic stem cells: long term stability, absence of senescence and a potential cell source for neural replacement. *Neuroscience* **145,** 1348–58 (2007).

133. Masutomi, K. *et al.* Telomerase maintains telomere structure in normal human cells. *Cell* **114,** 241–53 (2003).

134. Chen, Q. M., Prowse, K. R., Tu, V. C., Purdom, S. & Linskens, M. H. Uncoupling the senescent phenotype from telomere shortening in hydrogen peroxide-treated fibroblasts. *Exp. Cell Res.* **265,** 294–303 (2001).

135. Parrinello, S. *et al.* Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts. *Nat. Cell Biol.* **5,** 741–7 (2003).

136. Di Leonardo, A., Linke, S. P., Clarkin, K. & Wahl, G. M. DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts. *Genes Dev.* **8,** 2540–51 (1994).

137. Stein, G. H., Drullinger, L. F., Soulard, A. & Dulić, V. Differential roles for cyclin-dependent kinase inhibitors p21 and p16 in the mechanisms of senescence and differentiation in human fibroblasts. *Mol. Cell. Biol.* **19,** 2109–17 (1999).

138. Jacobs, J. J. L. & de Lange, T. Significant role for p16INK4a in p53-independent telomere-directed senescence. *Curr. Biol.* **14,** 2302–8 (2004).

139. Bakkenist, C. J. & Kastan, M. B. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* **421,** 499–506 (2003).

140. Roninson, I. B. Tumor cell senescence in cancer treatment. *Cancer Res.* **63,** 2705–15 (2003).

141. Von Zglinicki, T. Oxidative stress shortens telomeres. *Trends Biochem. Sci.* **27,** 339–44 (2002).

142. Narita, M. *et al.* Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113,** 703–16 (2003).

143. Kim, W. Y. & Sharpless, N. E. The regulation of INK4/ARF in cancer and aging. *Cell* **127,** 265–75 (2006).

144. Munro, J., Barr, N. I., Ireland, H., Morrison, V. & Parkinson, E. K. Histone deacetylase inhibitors induce a senescence-like state in human cells by a p16-

dependent mechanism that is independent of a mitotic clock. *Exp. Cell Res.* **295,** 525–38 (2004).

145. Taddei, A., Maison, C., Roche, D. & Almouzni, G. Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat. Cell Biol.* **3,** 114–20 (2001).

146. Bandyopadhyay, D. *et al.* Down-regulation of p300/CBP histone acetyltransferase activates a senescence checkpoint in human melanocytes. *Cancer Res.* **62,** 6231–9 (2002).

147. Minucci, S. & Pelicci, P. G. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat. Rev. Cancer* **6,** 38–51 (2006).

148. Itahana, K. *et al.* Control of the replicative life span of human fibroblasts by p16 and the polycomb protein Bmi-1. *Mol. Cell. Biol.* **23,** 389–401 (2003).

149. Krishnamurthy, J. *et al.* Ink4a/Arf expression is a biomarker of aging. *J. Clin. Invest.* **114,** 1299–307 (2004).

150. Zindy, F., Quelle, D. E., Roussel, M. F. & Sherr, C. J. Expression of the p16INK4a tumor suppressor versus other INK4 family members during mouse development and aging. *Oncogene* **15,** 203–11 (1997).

151. Janzen, V. *et al.* Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a. *Nature* **443,** 421–6 (2006).

152. Jacobs, J. J., Kieboom, K., Marino, S., DePinho, R. A. & van Lohuizen, M. The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397,** 164–8 (1999).

153. Freund, A., Patil, C. K. & Campisi, J. p38MAPK is a novel DNA damage response-independent regulator of the senescence-associated secretory phenotype. *EMBO J.* **30,** 1536–48 (2011).

154. Beauséjour, C. M. *et al.* Reversal of human cellular senescence: roles of the p53 and p16 pathways. *EMBO J.* **22,** 4212–22 (2003).

155. Narita, M. Cellular senescence and chromatin organisation. *Br. J. Cancer* **96,** 686–91 (2007).

156. Campisi, J. & d'Adda di Fagagna, F. Cellular senescence: when bad things happen to good cells. *Nat. Rev. Mol. Cell Biol.* **8,** 729–40 (2007).

157. Loeb, L. A. Cancer cells exhibit a mutator phenotype. *Adv. Cancer Res.* **72,** 25–56 (1998).

158. Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88,** 323–31 (1997).

159. Lengauer, C., Kinzler, K. W. & Vogelstein, B. Genetic instabilities in human cancers. *Nature* **396,** 643–9 (1998).

160. Bartkova, J. *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444,** 633–7 (2006).

161. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88,** 593–602 (1997).

162. Lin, A. W. *et al.* Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. *Genes Dev.* **12,** 3008–19 (1998).

163. Wajapeyee, N., Serra, R. W., Zhu, X., Mahalingam, M. & Green, M. R. Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. *Cell* **132,** 363–74 (2008).

164. Michaloglou, C. *et al.* BRAFE600-associated senescence-like cell cycle arrest of human naevi. *Nature* **436,** 720–4 (2005).

165. Dimri, G. P., Itahana, K., Acosta, M. & Campisi, J. Regulation of a senescence checkpoint response by the E2F1 transcription factor and p14(ARF) tumor suppressor. *Mol. Cell. Biol.* **20,** 273–85 (2000).

166. TODARO, G. J. & GREEN, H. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J. Cell Biol.* **17,** 299–313 (1963).

167. Bryan, T. M. & Cech, T. R. Telomerase and the maintenance of chromosome ends. *Curr. Opin. Cell Biol.* **11,** 318–24 (1999).

168. Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S. & Reddel, R. R. Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* **14,** 4240–8 (1995).

169. Woo, R. A. & Poon, R. Y. C. Activated oncogenes promote and cooperate with chromosomal instability for neoplastic transformation. *Genes Dev.* **18,** 1317–30 (2004).

170. Bringold, F. & Serrano, M. Tumor suppressors and oncogenes in cellular senescence. *Exp. Gerontol.* **35,** 317–29 (2000).

171. Itahana, K., Dimri, G. & Campisi, J. Regulation of cellular senescence by p53. *Eur. J. Biochem.* **268,** 2784–91 (2001).

172. Tahara, H., Sato, E., Noda, A. & Ide, T. Increase in expression level of p21sdi1/cip1/waf1 with increasing division age in both normal and SV40-transformed human fibroblasts. *Oncogene* **10,** 835–40 (1995).

173. Weinberg, R. A. The retinoblastoma protein and cell cycle control. *Cell* **81,** 323–30 (1995).

174. Pietenpol, J. A., Holt, J. T., Stein, R. W. & Moses, H. L. Transforming growth factor beta 1 suppression of c-myc gene transcription: role in inhibition of keratinocyte proliferation. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 3758–62 (1990).

175. Sherr, C. J. & Roberts, J. M. Inhibitors of mammalian G1 cyclin-dependent kinases. *Genes Dev.* **9,** 1149–63 (1995).

176. Di Micco, R. *et al.* Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* **444,** 638–42 (2006).

177. Chen, Z. *et al.* Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* **436,** 725–30 (2005).

178. Dhomen, N. *et al.* Oncogenic Braf induces melanocyte senescence and melanoma in mice. *Cancer Cell* **15,** 294–303 (2009).

179. Braig, M. *et al.* Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* **436,** 660–5 (2005).

180. Lazzerini Denchi, E., Attwooll, C., Pasini, D. & Helin, K. Deregulated E2F activity induces hyperplasia and senescence-like features in the mouse pituitary gland. *Mol. Cell. Biol.* **25,** 2660–72 (2005).

181. Hahn, W. C. & Weinberg, R. A. Rules for making human tumor cells. *N. Engl. J. Med.* **347,** 1593–603 (2002).

182. Ghebranious, N. & Donehower, L. A. Mouse models in tumor suppression. *Oncogene* **17,** 3385–400 (1998).

183. Boulanger, C. A. & Smith, G. H. Reducing mammary cancer risk through premature stem cell senescence. *Oncogene* **20,** 2264–72 (2001).

184. Collado, M. *et al.* Tumour biology: senescence in premalignant tumours. *Nature* **436,** 642 (2005).

185. Yang, G. *et al.* The chemokine growth-regulated oncogene 1 (Gro-1) links RAS signaling to the senescence of stromal fibroblasts and ovarian tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 16472–7 (2006).

186. Kortlever, R. M., Higgins, P. J. & Bernards, R. Plasminogen activator inhibitor-1 is a critical downstream target of p53 in the induction of replicative senescence. *Nat. Cell Biol.* **8,** 877–84 (2006).

187. McElhaney, J. E. & Effros, R. B. Immunosenescence: what does it mean to health outcomes in older adults? *Curr. Opin. Immunol.* **21,** 418–24 (2009).

188. Jun, J.-I. & Lau, L. F. The matricellular protein CCN1 induces fibroblast senescence and restricts fibrosis in cutaneous wound healing. *Nat. Cell Biol.* **12,** 676–685 (2010).

189. Gilbert, L. A. & Hemann, M. T. DNA damage-mediated induction of a chemoresistant niche. *Cell* **143,** 355–66 (2010).

190. Sun, Y. *et al.* Treatment-induced damage to the tumor microenvironment promotes prostate cancer therapy resistance through WNT16B. *Nat. Med.* **18,** 1359–68 (2012).

191. Ivy, J. M., Klar, A. J. & Hicks, J. B. Cloning and characterization of four SIR genes of Saccharomyces cerevisiae. *Mol. Cell. Biol.* **6,** 688–702 (1986).

192. Kaeberlein, M., McVey, M. & Guarente, L. The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms. *Genes Dev.* **13,** 2570–80 (1999).

193. Chen, D. & Guarente, L. SIR2: a potential target for calorie restriction mimetics. *Trends Mol. Med.* **13,** 64–71 (2007).

194. Dokmanovic, M., Clarke, C. & Marks, P. A. Histone deacetylase inhibitors: overview and perspectives. *Mol. Cancer Res.* **5,** 981–9 (2007).

195. Haigis, M. C. & Sinclair, D. A. Mammalian sirtuins: biological insights and disease relevance. *Annu. Rev. Pathol.* **5,** 253–95 (2010).

196. Sebastian, C., Satterstrom, F. K., Haigis, M. C. & Mostoslavsky, R. From Sirtuin Biology to Human Diseases: An Update. *J. Biol. Chem.* **287,** 42444–42452 (2012).

197. D'Mello, S. R. Histone deacetylases as targets for the treatment of human neurodegenerative diseases. *Drug News Perspect.* **22,** 513–24 (2009).

198. Cantó, C. & Auwerx, J. Caloric restriction, SIRT1 and longevity. *Trends Endocrinol. Metab.* **20,** 325–31 (2009).

199. Koubova, J. & Guarente, L. How does calorie restriction work? *Genes Dev.* **17,** 313–21 (2003).

200. Fontana, L., Partridge, L. & Longo, V. D. Extending healthy life span--from yeast to humans. *Science* **328,** 321–6 (2010).

201. Langley, E. *et al.* Human SIR2 deacetylates p53 and antagonizes PML/p53-induced cellular senescence. *EMBO J.* **21,** 2383–96 (2002).

202. Picard, F. *et al.* Sirt1 promotes fat mobilization in white adipocytes by repressing PPAR-gamma. *Nature* **429,** 771–6 (2004).

203. Van der Horst, A. *et al.* FOXO4 is acetylated upon peroxide stress and deacetylated by the longevity protein hSir2(SIRT1). *J. Biol. Chem.* **279,** 28873–9 (2004).

204. Cohen, H. Y. *et al.* Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Mol. Cell* **13,** 627–38 (2004).

205. Liu, T., Liu, P. Y. & Marshall, G. M. The critical role of the class III histone deacetylase SIRT1 in cancer. *Cancer Res.* **69,** 1702–5 (2009).

206. Chu, F., Chou, P. M., Zheng, X., Mirkin, B. L. & Rebbaa, A. Control of multidrug resistance gene mdr1 and cancer resistance to chemotherapy by the longevity gene sirt1. *Cancer Res.* **65,** 10183–7 (2005).

207. Liu, G. *et al.* Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in glioblastoma. *Mol. Cancer* **5,** 67 (2006).

208. Wang, R.-H. *et al.* Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell* **14,** 312–23 (2008).

209. Zhao, W. *et al.* Negative regulation of the deacetylase SIRT1 by DBC1. *Nature* **451,** 587–90 (2008).

210. Chen, W. Y. *et al.* Tumor suppressor HIC1 directly regulates SIRT1 to modulate p53-dependent DNA-damage responses. *Cell* **123,** 437–48 (2005).

211. Heltweg, B. *et al.* Antitumor activity of a small-molecule inhibitor of human silent information regulator 2 enzymes. *Cancer Res.* **66,** 4368–77 (2006).

212. Ota, H. *et al.* Sirt1 inhibitor, Sirtinol, induces senescence-like growth arrest with attenuated Ras-MAPK signaling in human cancer cells. *Oncogene* **25,** 176–85 (2006).

213. Suzuki, K. *et al.* SRT1720, a SIRT1 activator, promotes tumor cell migration, and lung metastasis of breast cancer in mice. *Oncol. Rep.* **27,** 1726–32 (2012).

214. Firestein, R. *et al.* The SIRT1 deacetylase suppresses intestinal tumorigenesis and colon cancer growth. *PLoS One* **3,** e2020 (2008).

215. Heltweg, B. *et al.* Antitumor activity of a small-molecule inhibitor of human silent information regulator 2 enzymes. *Cancer Res.* **66,** 4368–77 (2006).

216. Huang, J. *et al.* SIRT1 overexpression antagonizes cellular senescence with activated ERK/S6k1 signaling in human diploid fibroblasts. *PLoS One* **3,** e1710 (2008).

217. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–54 (1976).

218. Timothy V. Updyke, S. C. E. System for pH-neutral stable electrophoresis gel. In: United States Patent 6, 162, 338. 2000 (2011).

219. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–70 (2003).

220. Moore, S. On the Determination of Cystine as Cysteic Acid. *J. Biol. Chem.* **238,** 235–237 (1963).

221. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Anal. Chem.* **75,** 663–670 (2003).

222. Spellman, D. S., Deinhardt, K., Darie, C. C., Chao, M. V & Neubert, T. A. Stable isotopic labeling by amino acids in cultured primary neurons: application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. *Mol. Cell. Proteomics* **7,** 1067–76 (2008).

223. Sasaki, T. *et al.* Phosphorylation regulates SIRT1 function. *PLoS One* **3,** e4020 (2008).

224. Bai, B. *et al.* Cyclin-dependent kinase 5-mediated hyperphosphorylation of sirtuin-1 contributes to the development of endothelial senescence and atherosclerosis. *Circulation* **126,** 729–40 (2012).

225. Zhang, B. *et al.* SIRT3 overexpression antagonizes high glucose accelerated cellular senescence in human diploid fibroblasts via the SIRT3-FOXO1 signaling pathway. *Age (Dordr).* **35,** 2237–53 (2013).

226. Wang, X.-Q. *et al.* Decreased SIRT3 in aged human mesenchymal stromal/stem cells increases cellular susceptibility to oxidative stress. *J. Cell. Mol. Med.* **18,** 2298–310 (2014).

227. Nagai, K. *et al.* Depletion of SIRT6 causes cellular senescence, DNA damage, and telomere dysfunction in human chondrocytes. *Osteoarthritis Cartilage* **23,** 1412–20 (2015).

228. Liu, J.-P. & Chen, R. Stressed SIRT7: facing a crossroad of senescence and immortality. *Clin. Exp. Pharmacol. Physiol.* **42,** 567–9 (2015).

229. Dai, J. *et al.* Identification of degradation products formed during performic oxidation of peptides and proteins by high-performance liquid chromatography with matrix-assisted laser desorption/ionization and tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **19,** 1130–8 (2005).

230. Chowdhury, S. K. *et al.* Mass spectrometric identification of amino acid transformations during oxidation of peptides and proteins: modifications of methionine and tyrosine. *Anal. Chem.* **67,** 390–8 (1995).

231. Zougman, A., Selby, P. J. & Banks, R. E. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14,** 1006–0 (2014).

232. Collado, M. & Serrano, M. Senescence in tumours: evidence from mice and humans. *Nat. Rev. Cancer* **10,** 51–7 (2010).

233. Nardella, C., Clohessy, J. G., Alimonti, A. & Pandolfi, P. P. Pro-senescence therapy for cancer treatment. *Nat. Rev. Cancer* **11,** 503–11 (2011).

234. Acosta, J. C. & Gil, J. Senescence: a new weapon for cancer therapy. *Trends Cell Biol.* **22,** 211–9 (2012).

235. Ewald, J. A., Desotelle, J. A., Wilding, G. & Jarrard, D. F. Therapy-induced senescence in cancer. *J. Natl. Cancer Inst.* **102,** 1536–46 (2010).

236. Yeh, C.-S. *et al.* Fatty acid metabolism pathway play an important role in carcinogenesis of human colorectal cancers by Microarray-Bioinformatics analysis. *Cancer Lett.* **233,** 297–308 (2006).

237. Hopperton, K. E., Duncan, R. E., Bazinet, R. P. & Archer, M. C. Fatty acid synthase plays a role in cancer metabolism beyond providing fatty acids for phospholipid synthesis or sustaining elevations in glycolytic activity. *Exp. Cell Res.* **320,** 302–10 (2014).

238. Hess, D. & Igal, R. A. Genistein downregulates de novo lipid synthesis and impairs cell proliferation in human lung cancer cells. *Exp. Biol. Med. (Maywood).* **236,** 707–13 (2011).

239. Liu, Z. L. *et al.* Inhibition of fatty acid synthase suppresses osteosarcoma cell invasion and migration via downregulation of the PI3K/Akt signaling pathway in vitro. *Mol. Med. Rep.* **7,** 608–12 (2013).

240. Migita, T. *et al.* Fatty acid synthase: a metabolic enzyme and candidate oncogene in prostate cancer. *J. Natl. Cancer Inst.* **101,** 519–32 (2009).

241. Zecchin, K. G. *et al.* Inhibition of fatty acid synthase in melanoma cells activates the intrinsic pathway of apoptosis. *Lab. Invest.* **91,** 232–40 (2011).

242. Orita, H., Coulter, J., Tully, E., Kuhajda, F. P. & Gabrielson, E. Inhibiting fatty acid synthase for chemoprevention of chemically induced lung tumors. *Clin. Cancer Res.* **14,** 2458–64 (2008).

243. Flavin, R., Peluso, S., Nguyen, P. L. & Loda, M. Fatty acid synthase as a potential therapeutic target in cancer. *Future Oncol.* **6,** 551–62 (2010).

244. Ford, J. H. Saturated fatty acid metabolism is key link between cell division, cancer, and senescence in cellular and whole organism aging. *Age (Dordr).* **32,** 231–7 (2010).

245. Veigel, D. *et al.* Fatty acid synthase is a metabolic marker of cell proliferation rather than malignancy in ovarian cancer and its precursor cells. *Int. J. Cancer* **136,** 2078–90 (2015).

246. Snel, B., Lehmann, G., Bork, P. & Huynen, M. A. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **28,** 3442–4 (2000).

247. Laberge, R.-M. *et al.* MTOR regulates the pro-tumorigenic senescence-associated secretory phenotype by promoting IL1A translation. *Nat. Cell Biol.* **17,** 1049–1061 (2015).

248. Kuo, L. J. & Yang, L.-X. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo* **22,** 305–9

249. Young, A. R. J. *et al.* Autophagy mediates the mitotic senescence transition. *Genes Dev.* **23,** 798–803 (2009).

250. Davis, T. A., Loos, B. & Engelbrecht, A.-M. AHNAK: the giant jack of all trades. *Cell. Signal.* **26,** 2683–93 (2014).

251. Lee, I. H. *et al.* Ahnak functions as a tumor suppressor via modulation of TGFβ/Smad signaling pathway. *Oncogene* **33,** 4675–84 (2014).

252. Hong, M. L. W., Jiang, N., Gopinath, S. & Chew, F. T. Proteomics technology and therapeutics. *Clin. Exp. Pharmacol. Physiol.* **33,** 563–8 (2006).

253. Mass Spectrometry | Max Planck Institute of Biochemistry. at <http://www.biochem.mpg.de/221765/Mass_spectrometry>

254. Roepstorff, P. & Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11,** 601 (1984).

255. Quantitative Proteomics | biOMICS. at <http://biomics.group.shef.ac.uk/research-portfolio/quantitative-proteomics/>

256. Huber, L. A. Is proteomics heading in the wrong direction? *Nat. Rev. Mol. Cell Biol.* **4,** 74–80 (2003).